NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

BULK METEOROLOGICAL PARAMETERS FOR DIAGNOSING CLOUDINESS IN THE STOCHASTIC CLOUD FORECAST MODEL
by
Ryan N. Leach
March 2006

Thesis Advisor:	Karl D. Pfeiffer
Second Reader:	Carlyle H. Wash

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** March 2006	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE: Bulk Meteorological Parameters for Diagnosing Cloudiness in the Stochastic Cloud Forecast Model.

5. FUNDING NUMBERS

6. AUTHOR(S) Ryan N. Leach, 2Lt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	**12b. DISTRIBUTION CODE**
Approved for public release; distribution is unlimited	A

13. ABSTRACT (maximum 200 words)

The three dimensional distribution of clouds is of great interest to the Air Force, and to the aviation community in general. The Stochastic Cloud Forecast Model (SCFM) is a novel, global cloud model currently operated at the Air Force Weather Agency (AFWA) which diagnoses cloud cover statistically using a minimal set of predictors from global numerical forecasts. Currently the four predictors are pressure, temperature, vertical velocity, and relative humidity. In this thesis, 330 sets of predictors are compared in the SCFM-R, a research version of the model programmed for this thesis. There are some differences in the SCFM and the SCFM-R that yield important information. It is found that the SCFM is very sensitive to how cloud cover in the boundary layer is diagnosed. An analysis of the diagnosis method used to initialize the model revealed a bias for over-diagnosing cloud at lower levels and under-diagnosing cloud at upper levels. Also, it is recommended that AFWA consider exchanging temperature for another predictor more related to moisture, such as cloud water, and that relative humidity is included as relative humidity to the fourth power. Other recommendations include improving the method for diagnosing cloud cover in the boundary layer and improving the model initial condition.

14. SUBJECT TERMS Clouds, Cloud Forecasting, Stochastic Cloud Model, Cloud Diagnosis, Stochastic Cloud Forecast Model

15. NUMBER OF PAGES 75

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

BULK METEOROLOGICAL PARAMETERS FOR DIAGNOSING CLOUDINESS IN THE STOCHASTIC CLOUD FORECAST MODEL

Ryan N. Leach
Second Lieutenant, United States Air Force
B.S., Washington State University, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 2006

Author: Ryan N. Leach

Approved by: Karl D. Pfeiffer, Lt Col, USAF
Thesis Advisor

Carlyle H. Wash
Second Reader

Philip A. Durkee
Chairman, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The three dimensional distribution of clouds is of great interest to the Air Force, and to the aviation community in general. The Stochastic Cloud Forecast Model (SCFM) is a novel, global cloud model currently operated at the Air Force Weather Agency (AFWA) which diagnoses cloud cover statistically using a minimal set of predictors from global numerical forecasts. Currently the four predictors are pressure, temperature, vertical velocity, and relative humidity. In this thesis, 330 sets of predictors are compared in the SCFM-R, a research version of the model programmed for this thesis. There are some differences in the SCFM and the SCFM-R that yield important information. It is found that the SCFM is very sensitive to how cloud cover in the boundary layer is diagnosed. An analysis of the diagnosis method used to initialize the model revealed a bias for over-diagnosing cloud at lower levels and under-diagnosing cloud at upper levels. Also, it is recommended that AFWA consider exchanging temperature for another predictor more related to moisture, such as cloud water, and that relative humidity is included as relative humidity to the fourth power. Other recommendations include improving the method for diagnosing cloud cover in the boundary layer and improving the model initial condition.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I greatly appreciate the many people who have given insight, support, and guidance during the research and writing of this thesis. First my advisor, Lt. Col Karl Pfeiffer, for his help with data, for connecting me with the right people at AFWA, and most of all for his enthusiasm for my work. Professor Carlyle Wash consistently provided a fresh set of eyes and ears as well as a non-military perspective that helped me keep the big picture in sight. Thanks are also due to Dr. David McDonald at AFWA/DNXM for his rapid answers to all my questions about all things SCFM, and to Bob Creasey for his efforts to appease my many computer resource needs. Finally I would like to thank my wife, Kari, for her patience, and for listening to me proofread my thesis aloud.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. THE NEED FOR CLOUD FORECASTING CAPABILITIES

In the civilian sector, as well as the Department of Defense (DOD), the presence of clouds, especially low clouds, causes the closure of runways, and limits flights based on pilot skill and aircraft capabilities. Norquist et al (1994; 1997; 1999; 2000) and Pukall (1998) discuss the importance of clouds to Air Force operations. The main themes include the restricted visibility caused by clouds, and their interference with electromagnetic systems. Reduced visibility interferes with any operation requiring air crew to make visual contact with an objective, such as in close air support, air drop, aerial refueling, and air-to-air combat. Interference with electromagnetic systems severely limits aerial and satellite-based reconnaissance and some weapons systems.

Cloud impacts are not only limited to the visible spectrum. Whether or not a wavelength of light is scattered by spherical cloud drops is primarily dictated by the size parameter,

$$\chi \equiv \frac{2\pi r}{\lambda} \tag{1}$$

where r is the radius of the cloud drop and λ is the wavelength of the light. If the size parameter is less than 10^{-3}, then scattering is negligible (Kidder and Vonder Harr, 1995). The typical cloud droplet size is 10 microns; however, clouds often have a broad spectra of drop sizes that can range between 5 microns and 100 microns, with 50 microns considered a large drop (Rogers and Yau, 1989). Using these criteria, wavelengths longer than 10^5 microns are needed to avoid scattering by clouds. Therefore, all wavelengths useable for passive sensing of the surface from solar radiation (visible light) to Earth's blackbody radiation (infrared and some microwave) are significantly scattered by clouds. Meteorologists use this to their advantage to observe clouds and the atmosphere. However, this is a serious impediment to anyone attempting to observe surface or near-surface properties from above. As a result, knowledge of the four-dimensional distribution of cloud is crucial to effective planning for reconnaissance and weapons systems employment.

B. THE CHALLENGES FOR CLOUD MODELS

Physically-based cloud models have several challenges. One of the most demanding is resolution. While it is theoretically possible to run a global model with cloud scale resolution, practically, there is not enough computer power to operationally run such a model in a timely fashion. For instance, it may take more than 48 h to make a 48 h forecast. Furthermore, there currently is no observing system with both the horizontal and vertical resolution needed to accurately initialize a model, especially on the global scale. Finally, certain properties of the atmosphere important to clouds, such as cloud condensation nuclei and aerosols, are not measured and modeled operationally on a global scale.

Stochastic cloud models may represent the best compromise between available technology and operational requirements. The statistical models discussed in this study all diagnose cloud from the forecast fields of the underlying NWP model; the forecasting is done in the underlying NWP model, and the statistical models infer, or diagnose, cloudiness from these forecasts. It is preferred that the forecast fields from the global NWP model relate directly to the physics of cloud formation, transport and sustainment, or decay; however, an indirect relationship may yield enough information to diagnose cloud. If predictors from the underlying NWP model have a physically based relationship with clouds, the modeler can have confidence that the model is emulating real cloud physics.

Currently, the most detailed and accessible information about the atmosphere on a global scale comes from model analysis and forecast fields of bulk meteorological parameters. There is also an abundance of satellite data available; unfortunately, this data usually offers poor vertical resolution. Currently, the United States Air Force (USAF) is producing a world wide nephanalysis, called the World Wide Merged Cloud Analysis (WWMCA), from multiple channels of data from multiple polar orbiting and geostationary satellites (HQ AFWA/DNXM, 2005). Several statistical cloud models have been based in finding statistical relationships between the clouds depicted in the WWMCA, or its predecessors, and the bulk meteorological parameters from NWP models (McDonald, 2003; Norquist et al., 1994; Norquist et al., 1997; Norquist, 1999;

Norquist, 2000; Pukall, 1998). The fitness of any given bulk parameter, or set of bulk parameters, for prediction of clouds may be highly dependent on the statistical model, the underlying NWP model, their physical relationship to clouds, and the environment of the clouds (e.g. tropical convective clouds versus clouds in a mid-latitude cyclone). The goal of this research has been to evaluate the fitness of many sets of bulk meteorological parameters for use with the Stochastic Cloud Forecast Model (SCFM), a global model currently employed by the USAF, and providing guidance for improving the SCFM.

The Air Force and others have been analyzing and forecasting cloudiness for many years. This will be discussed in detail in Chapter II. One of the more recent attempts to forecast clouds by the Air Force is the SCFM. The SCFM is a statistical model that globally forecasts cloudiness using a look up table of four predictors from an NWP model and the two-parameter beta distribution (McDonald, 2003). Chapter III explains the details of the operational version of the SCFM, the SCFM16, and the details of the research version used for this study, the SCFM-R. Chapter III also discusses how different sets of predictors will be assessed using the SCFM-R. The main thrust of this study is to find the best set of four predictors practically possible for the SCFM; these results are presented in Chapter IV. Chapter V discusses model consistency with other statistical cloud modeling efforts, other possible model configurations, and verification issues. Conclusions and recommendations are given in Chapter VI.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. DEFINING CLOUD

When modeling clouds, it is important to define what exactly is being modeled, and depending on the perspective of the modeler, "cloud" can have different meanings. From a meteorology perspective, cloud refers to the microscopic water drops suspended in the atmosphere; in that sense clouds are assessed by their scattering of electromagnetic radiation. An aviator, however, may think of clouds in terms of an obstruction to visibility. These two perspectives lead to two categories of cloud definitions in models and measurements: physical definitions and definitions based on impact to daily operations.

When treating clouds in NWP models, meteorologists attempt to model the physical characteristics of clouds. Some physical characteristics of cloud that could be modeled include the optical depth of the clouds, the mean droplet size, the droplet size distribution, or the number density of drops. More often cloud water and/or ice is modeled as the mass of liquid or ice per kilogram of air. In some models, such as the Mesoscale Model 5 (MM5) (Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research, 2002), there are complex models of clouds that include the type of precipitation that is produced, such as rain, snow, hail, rime, or freezing rain. While these models are very detailed about the physical properties of clouds, they tell users very little about how clouds specifically will impact operations.

Normally, clouds impact people by reducing visibility, obstructing a line of sight, or reducing the amount of sunlight received during the day. Jensenius (1988) measured this type of impact by not measuring clouds at all, but measuring how much time during the day that the disk of the sun was completely unobstructed by clouds. A more common definition of cloud is used by the National Weather Service in their model output statistics (MOS) for clouds (Weiss, 2001), which was developed and tested by Glahn and Lowry (1972), and again by Glahn and Carter (1976). These researchers define cloud by the fraction of the sky that is obscured from the ground by cloud. A similar measure of cloud, the fraction of the ground obstructed by cloud as viewed from space has been in

common use by the Air Force for many years too (McDonald, 2003; Norquist et al., 1994; Norquist et al., 1997; Norquist, 1999; Norquist, 2000; Pukall, 1998; HQ AFWA/DNXM, 2005; Whiton et al., 1981). Arguably, these two definitions should be equivalent, since it should not matter from which direction you are observing the clouds. Many have not addressed the issue of whether the clouds are being observed from the ground or space, only that there is a fractional cloudiness (Kvamsto, 1991; Kvamsto, 1993; Mocko and Cotton, 1995). Since cloud observations in the following refer to satellite cloud observations, hereafter in this thesis cloud measurements will be considered from the point of view of space, or at least from the point of view of nadir.

B. CLOUD ANALYSIS

For any model, verification data is needed, and in the case of statistical models “ground truth” data is also needed to train the model. Training the model is the process of adjusting model parameters to minimize the model error. For the SCFM, the predictand is cloud cover, and the only current measurement of cloud cover on a global scale is the WWMCA from the Cloud Depiction and Forecast System II (CDFS-II) operated by the Air Force (HQ AFWA/DNXM, 2005). For complete documentation of the WWMCA, see the Algorithm Description for the Cloud Depiction and Forecast System II (HQ AFWA/DNXM, 2005) from the Air Force Weather Agency (AFWA) at Offut AFB, Nebraska. Below, the important features of the WWMCA for the SCFM are summarized from the Algorithm Description for the Cloud Depiction and Forecast System II.

Data for the WWMCA is taken from both geostationary and polar orbiting satellites. This includes all currently operating GOES and MeteoSat satellites for geostationary orbits, and the NOAA polar orbiting satellites with the AVHRR sensor package and the DMSP satellite with the OLS sensor package for polar orbiters. The geostationary satellites provide good time resolution, while the polar orbiters provide good spatial resolution and polar coverage that geostationary satellites cannot measure.

After ingesting the heterogeneous satellite data, the first task of the CDFS-II is to perform cloud discrimination for each satellite. This is done with several different techniques. One test, commonly used on visible (VIS) and infrared (IR) channels, is a threshold test. In threshold tests, the radiance for IR channels or the reflectance for VIS

channels is compared to an expected threshold. If the radiance or reflectance is beyond the threshold, then that pixel is deemed as cloudy. Expected values are calculated from climatology, previous measurements, NWP model temperatures of the ground, and empirical relationships. On satellites with multiple IR and VIS channels, like most of the current polar orbiters, ratios of radiance or reflectivity can be used to discern between clouds and snow or dust. Due to the good time resolution of the geostationary satellites, temporal tests can be done on their data. Temporal tests detect sudden changes in reflectance or radiance, which in turn implies cloud advection. On each pixel of each satellite image, as many tests as possible are run, and the level of agreement between them is used to assign a confidence level for that pixel. It is important to note for each satellite pixel, cloud discrimination within the CDFS-II is a binary process: either there is cloud, or there is not cloud.

Once the cloud discrimination has been completed, each satellite's data is transferred to two 24 km, 16^{th}-mesh grids, one for the northern hemisphere and one for the southern hemisphere, and merged. The 16^{th} mesh grid is a polar stereographic map projection; the m^{th} mesh grid system is a system that has been commonly used at AFWA for over 25 years (Hoke, Dr. James E. et al., 1981). When a satellite's data is transferred to a 16^{th}-mesh grid, more than one pixel from the satellite image will correspond to each 16^{th}-mesh grid-box. Cloud cover for a grid-box is then calculated as the percentage of cloudy pixels in that grid-box. The merging process consists of choosing which satellite's data to use for each grid-box. Factors taken into consideration include the confidence level computed in the cloud discrimination step and age of the data from each satellite.

Finally on the final 16^{th}-mesh grid, various statistical routines are employed to determine cloud layers and types. Cloud tops are inferred from comparisons with model temperature soundings and the radiances of IR data, and cloud bases are calculated statistically with empirical relationships. These routines are very complex, and beyond the scope of this thesis. The final product, the WWMCA, has the following fields: one field indicating the age of the data at the WWMCA valid time, one field of the total cloud amount, four fields of layer cloud amounts, four layers of cloud tops, four layers of cloud bases, and four layers of cloud types. The SCFM uses the total cloud record, and none of

the cloud layer information because this information is considered to be unreliable by some investigators (McDonald, 2005).

C. CLOUD MODELS

The conceptual definitions of cloud, physical definitions and impact related definitions each correlate to a different framework for modeling clouds. Physical cloud models include either direct or parameterized implementations of clouds. In fact, most, if not all, of the currently running operational NWP models include clouds in some way to account for their physical effects on the atmosphere (Kalnay, 2003). Statistical models take information about the physical aspects of clouds, or the atmosphere in general, from NWP models and other analyses and transform it into information about cloud impacts. For the purpose of this thesis, empirical models are grouped with statistical models, since the problem of determining the model parameters is generally statistical.

1. Physical Models

The most direct physical cloud models are models that integrate the primitive equations for the atmosphere with explicit treatment of moisture (Rogers and Yau, 1989). However, the spatial resolution required to do this is extremely high since cumulus clouds commonly occur at sizes of 2-5 km (Fujita, 1986). Currently, the Global Forecast System (GFS) operated by the National Center for Environmental Prediction (NCEP) has a resolution of 382 waves, or T382 (Campana et al., 2005). This works out to be almost 35 km at the equator, which is far coarser than cloud scale at a place on the globe where cumulus convection is the primary driving force in the atmosphere.

Most physical models used operationally employ parameterizations to account for the role of clouds in the atmosphere. However, since clouds can occur on scales smaller than the grid spacing of such models, these parameterizations model the average affect of clouds over the whole grid box (Kalnay, 2003). Some mesoscale models have very detailed microphysics included in the cloud parameterizations, such as the MM5 (Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research, 2002).

A simple type of physical cloud model is an advection model. Cloud advection models ingest a three dimensional nephanalysis, such as the WWMCA, and wind fields from an NWP model and move the clouds as if they were floating in the wind. The Air

Force has used different models that do exactly that since the late 1960's (Crum, Timothy D. ed., 1987). Research with these models has shown that they can have considerable skill for short time periods from 9-36 hours (Pukall, 1998). Hamill and Nehrkorn have attempted to do such a trajectory model using statistics to measure winds (Hamill and Nehrkorn, 1993). They used satellite images only, instead of working on a model grid, and cross correlation to derive winds from cloud motions. They would then forecast clouds by advecting them and assuming there would be no change in the winds. This forecast method proved to be more skillful than persistence alone; however, it had difficulty in regions with rapid cloud formation, such as convective regions. Difficulty also arose in calculating or measuring winds via interpolation or extrapolation in clear areas. In general, a weakness of cloud advection models is that they do not account for cloud formation and evaporation. However, the Air Force has attempted to parameterize these effects with some success (Crum, Timothy D. ed., 1987).

2. Statistical Models

In general, statistical cloud models are models that post process some form of physical NWP in order to extend the information from it, or transform the information in it into a form more relevant to the end user of the model. That is, statistical cloud models are not a substitute for integrating the primitive equations as done by physical NWP models, they are an extension of such models. Below, four important and predominant statistical cloud forecasting models, or classes of models, are discussed; this is by no means an exhaustive review.

a. Simple Empirical Models

Mocko and Cotton (1995) discuss several models for calculating cloud cover in the boundary layer ranging from simple schemes that only employ relative humidity to complex schemes that take into account the sub-grid variability of the underlying NWP model, which was the same for every scheme. They found that some of the simplest schemes based primarily on relative humidity performed the best. One of the schemes, proposed by Kvamsto (1991; 1993), is a simple linear relationship between relative humidity and cloud cover:

$$CC = \left(\frac{Rh - Rh_{00}}{Rh_s - Rh_{00}}\right) * 100 \qquad (2)$$

where CC is the cloud cover in percent, Rh is the relative humidity, Rh_{00} is a critical relative humidity below which no condensation occurs, and Rh_s is the relative humidity above which it is overcast. If $Rh < Rh_{00}$, then CC is zero, there is no cloud cover, and if $Rh > Rh_s$, then CC is 100%. An adaptaion of Kvamsto's CC equation (eq. 2) is used in the SCFM as an initial condition. Another simple cloud cover model that performed well in the tests by Mocko and Cotton was proposed by Sundqvist et al (1989). Sundqvist et al used a different functional form of the same ratio as equation 2 above. Mocko and Cotton note that the methods proposed by Sundqvist et al and Kvamsto have performed well with underlying NWP models with grid scales ranging from 5 to 80 km.

Further research on the schemes proposed by Kvamsto and Sundqvist et al was conducted by Pukall (1998), when he calibrated them for use in the MM5. Pukall found that when comparing these models used with the MM5 to a cloud advection model, both models outperformed the cloud advection scheme, but performed similar to each other. Cloud condensate from the MM5 was also used by Pukall to diagnose cloud. A column integrated cloud condensate approach was taken as well as a layer by layer approach for using cloud condensate to diagnose cloud cover. Thresholds, which Pukall determined empirically, were used to translate different levels of cloud condensate into cloud cover. The cloud condensate scheme performance was similar to the relative humidity schemes discussed above. Pukall noted that the most significant weakness for all of these methods for his tests was the underlying biases of the NWP model.

b. Model Output Statistics

The term Model Output Statistics (MOS) was coined in 1972 by Glahn and Lowry (1972). MOS can be used to relate any weather element not forecast well by physical NWP models to fields that the NWP model does forecast well, which is the goal of any statistical weather model in general. In the case of cloud cover, some models do not even directly forecast the desired quantity, so MOS is used (Carter and Glahn, 1976; Glahn and Lowry, 1972). In some cases, such as MOS output for the GFS by NCEP, predictors can be recent observations or soundings from a nearby station (Weiss, 2001). The basic statistical principle behind MOS is multiple linear regression.

In multiple linear regression, a single predictand is modeled as a linear combination of several predictors. For MOS, a different forecast equation is developed

for each forecast time, that is, a separate equation is developed for the 3, 6, 9, 12, etc., hour forecasts. This way MOS can also account for predominant model errors and biases. Predictors are chosen by stepwise screening regression. In this method, all of the available predictors are made available, and at each step the predictor that reduces the root mean square error (RMSE) the most on the dependent data set is added to the equation (Glahn and Lowry, 1972). If the reduction of the RMSE does not exceed a given threshold, then no more predictors are selected.

Currently, the GFS model at NCEP has MOS distributed for the 0000 and 1200Z forecast runs (Weiss, 2001). The number of predictors was not allowed to exceed 18 for cloud cover and cloud base predictions. Possible predictors for the GFS model and cloud cover consisted of relative humidity variables, wind components, wind speed, vertical velocity, relative vorticity, absolute vorticity advection, the lifted condensation level, relative humidity times vertical velocity, total precipitable water, three hour precipitation amount, and the K-index (Weiss, 2001).

MOS does not produce a general statistical model. Many MOS equations are developed for a single station (Glahn and Lowry, 1972), whereas others are developed for a region (Weiss, 2001). The "truth" data used to train MOS equations are local in-situ observations. Better performance of MOS occurs also when different equation sets are developed for different seasons (Weiss, 2001). MOS equations are developed once and applied thereafter to model output, without further training. This means that whenever major changes are made to the underlying NWP model, a whole new set of MOS equations has to be developed.

c. Diagnostic Cloud Forecast Model

The Diagnostic Cloud Forecast Model (DCFM) is a regional scale statistical cloud model operated by the Air Force, and is based on the work done by Norquist et al (1994; 1997; 1999; 2000). Norquist et al experimented with several types of statistical forecasting techniques, and finally chose a multiple linear discriminant approach that shares some similarities with MOS, but is much more general. The real world data used to train the DCFM is the WWMCA. Methods tested included multiple linear discriminant analysis (MDA), multiple linear regression (MLR), and MLR with regression estimation of event probabilities (REEP). REEP is applied to categorical

predictands; the regression equations return the probability of each category being the correct one. REEP is also used in MOS, when the predictand is categorical (Weiss, 2001). MDA was finally selected by Norquist et al as the method of choice.

However, Norquist et al (2000) noticed that just choosing model parameters that minimized the RMSE did not produce a realistic looking cloud forecast. The distribution of the forecast cloud field was very different from the distribution of the observed cloud field, and as a result, the cloud forecasts did not look realistic. To remedy this, Norquist et al developed their final model to output probabilities of a point falling into a cloudiness category. Then, instead of assigning the cloudiness value that was most probable to each point, cloudiness values were assigned so that each point had a cloudiness that was highly probable, but also ensured the overall distribution of the cloud field would match the observed cloud field distribution. This only produced a relatively minor increase in the RMSE, but it produced much more realistic cloud forecasts.

The DCFM works within a 10-day window prior to the forecast initialization time. Data from that 10-day period is used to train MDA equations much the same way MOS is done, with forward stepwise predictor selection. However, unlike MOS, the DCFM does not use previous observations, and the equations developed are general for the whole model grid. This way every grid point of a given grid has a cloud diagnosis, and a forecast field of cloud can be generated instead of a point forecast. The 10-day sliding window for data allows the DCFM to quickly adapt to changes in any underlying NWP model, as well as adapt to changes in the environment due to changing seasons. For every forecast made by the DCFM, new equations are developed; this means that different predictors can be chosen each time.

The DCFM's greatest strength is its ability to adapt quickly to new models. When major changes are made to a model, or a new model is desired for the underlying NWP forecast, the DCFM can be trained and ready to go 10 days after the switch is made. However, the DCFM has some significant weaknesses also, including linearity and cumbersome operation. Because the DCFM has to train so many equations so frequently, it is a very cumbersome and complicated model, which may be why it isn't used on global scales. The DCFM, like MOS, assumes linear relationships between the

predictors and predictand. However, from what is known about cloud physics, the physical processes that make, maintain, and dissipate clouds are highly nonlinear (Rogers and Yau, 1989).

d. Stochastic Cloud Forecast Model

The details of the SCFM are discussed in Chapter 3, however; a short description here will serve to place the SCFM into context with the other statistical models discussed. The SCFM was also developed by the Air Force, and has primarily been the work of McDonald (2003; 2005). Currently it uses four predictors: pressure, temperature, vertical velocity, and relative humidity. These predictors are discretized, or binned into categories, and used as indices in a look up table (LUT). To train the SCFM, cloud cover at each level of the model is inferred from the GFS analysis and the WWMCA, and statistics about the cloud amount, in percent, are stored in the cells of the LUT, based on the indices given by the predictors for that point. These statistics are then used to infer the distribution of cloud amount in that cell of the LUT. When a forecast is to be made for a point, the predictors at a point from a GFS forecast are used to find the corresponding cell in the LUT, and a random deviate is generated based on the distribution of cloud described by the statistics in that cell of the LUT. More detailed explanations of each of these steps will be given below in Chapter 3.

The advantages of the SCFM are its lack of linearity assumptions, and its small size. Linear models assume that a single point in predictor space corresponds to exactly a single output. However, the SCFM's creators realized that sub-grid scale variability was often important for clouds. So they only assumed that large-scale bulk conditions corresponded to a distribution of possible cloud amounts, hence the distributions in the cells of the LUT. Observations showed that a two-parameter beta distribution, which will be discussed below, was a good choice of a distribution to fit, which is the only other assumption in the model. By developing the distribution in each cell of the LUT independently, the linearity assumption used by the DCFM and MOS is removed; this means in general that no functional form has to be assumed, whether linear, or nonlinear.

D. RELATIONSHIPS OF SOME POSSIBLE PREDICTORS TO CLOUD

It is well known among meteorologists that clouds, or their absence, can yield much information about the current dynamical state of the atmosphere (Djuric, 1994; Kidder et al, 1995). Essentially, statistical cloud forecast models that diagnose cloud cover from an underlying NWP forecast are inverting this process to yield information about cloud cover from the dynamics of the atmosphere as modeled by NWP processes. Below, the predictors available for this study and their relationships to clouds are discussed.

1. Pressure, Geopotential, and Potential Temperature

Pressure, geopotential height, and potential temperature serve mainly as proxies for height; in fact, these parameters often serve as the vertical coordinate in NWP models (Kalnay, 2003). While height alone does not offer any information about cloud cover, it may be useful for breaking the atmosphere into regions. For instance, a relatively low temperature at 500 mb may indicate a cold pool or the passage of a short wave, which may be associated with clouds. Pressure and geopotential together could indicate the presence of a ridge/trough system, which in turn can be related to an increased/decreased probability for cloud cover. Height could be very effective when used with another predictor that varies greatly with height, such as vorticity. By separating out each level using pressure or geopotential height, the remaining predictors will be free to only diagnose cloud related conditions as they occur on that single level. For this reason, a proxy for height, such as pressure or geopotential, is hypothesized to be among the sets of predictors that perform well.

2. Temperature

Temperature alone does not imply an amount of cloud cover. However, it may be useful in combination with other predictors. Physically, temperature is directly related to cloud formation and sustainment through its role in determining the saturation vapor pressure of water in the atmosphere (Rogers and Yau, 1989). Temperature is also important for determining the phase of the clouds. Since temperature generally decreases with height, although not necessarily monotonically, it could also be used as a proxy for height. However, since pressure and geopotential are much better proxies for height and

temperature alone cannot be directly related to cloud cover, temperature is not hypothesized to be among the sets of predictors that perform well. However, it should be noted that AFWA is currently using temperature as a predictor successfully. This may be related to the fact that the initial condition used to initially diagnose cloud cover (see Chapter 3 below) is in fact an empirical relationship that depends solely on relative humidity and temperature.

3. Vertical Velocity

Vertical velocity is directly related to cloud cover. Any text on cloud physics, such as Rogers and Yau (1989), will discuss the importance of vertical motion to clouds. To review, as air rises it cools adiabatically until saturation is reached. After saturation is reached, condensation starts and cloud droplets form. Conversely, as air sinks it warms adiabatically and water drops evaporate dissipating the cloud. In fact, the creation or dissipation of clouds in satellite images is interpreted as vertical motion in the atmosphere (Djuric, 1994; Kidder and Vonder Harr, 1995). Therefore it is logical that vertical motion should be a good predictor of cloud cover, and it is expected to be among the predictors in the sets of predictors that perform well.

4. Absolute Vorticity and Relative Vorticity

Vorticity can be related to clouds indirectly through its association with weather patterns, such as mid-latitude cyclones and anticyclones. These weather patterns are generally related to cloud patterns; in the case of mid-latitude cyclones, the dynamic structure of the storm is related to clouds with conveyor belt theories. The relationship of vorticity to these patterns is described by the quasi-geostrophic equations. Fronts also have strong vorticity and cloud signatures (Bluestein, 1993; Carlson, 1998). Another example of a system with both strong cloud and vorticity signatures is tropical cyclones. Exploiting these indirect relationships will yield information about cloud cover. However, it should be noted that vorticity, in terms of vorticity advection, is related to vertical motion through the quasi-geostrophic ω equation (Bluestein, 1993; Carlson, 1998). This suggests that vertical velocity may be a more useful predictor, since it is closely related to vorticity and it is more directly related to clouds.

5. Relative Humidity and Relative Humidity^4

Relative humidity is the most intuitive predictor for cloud cover. If the atmosphere is too dry for the ambient temperature and pressure, then clouds cannot form. Likewise, if the atmosphere is too moist for the ambient temperature and pressure, then condensation will occur. In fact, the nucleation rate of cloud droplets and the growth rate of cloud droplets and ice crystals is strongly dependent on the relative humidity of the atmosphere (Rogers and Yau, 1989). However, there are other very important factors for cloud formation, sustainment, and decay to include: the presence of condensation nuclei, the actual amount of water vapor available (specific humidity), and the effects of solutes on cloud droplets (Rogers and Yau, 1989). Other studies have shown that the relative humidity at which condensation occurs in the atmosphere can strongly depend on temperature (Inoue and Kamahori, 2001; Kvamsto, 1991; Kvamsto, 1993; McDonald, 2003; Mocko and Cotton, 1995; Pukall, 1998). Nevertheless, relative humidity in some form is expected to be a very important predictor for this study. Reexpressing relative humidity (RH) as RH^4 may actually improve its performance. Most condensation occurs at high relative humidities and cloud drop nucleation is extremely sensitive to humidity (Rogers and Yau, 1989), and this reexpression condenses the low relative humidities into a few low categories, and expands the high relative humidities over more categories, allowing for more detail in the model at the sensitive relative humidities.

6. Cloud Water

Cloud water is the direct representation of clouds for thermodynamic and moisture budget purposes in NWP models (Kalnay, 2003). Cloud water can be used to infer cloud cover, as was attempted by Pukall (1998). Because of this very direct relationship to cloud cover, cloud water is expected to be a strong predictor of cloud cover. However, it should be noted that any weaknesses of the NWP model's parameterization of clouds will directly affect the cloud water in the model. As a result, it is expected that cloud water alone is not enough to accurately diagnose cloud cover.

7. Product of Relative Humidity and Pressure Vertical Velocity

Relative humidity and vertical velocity are both very directly related to cloud cover, and it is hypothesized that some combination of the two will also be a good predictor. Their product is a good choice. When their product is negative, upward

vertical motion is implied (in pressure coordinates), and when its magnitude is large moist air is implied. So large negative values indicate moist air rising, and thus cloud cover. Small positive values indicate dry air sinking, and thus clear skies. Others have used the product of vertical velocity and relative humidity with success (Weiss, 2001).

8. Other Possible Predictors and Motivation for Testing a Predictor

The number of possible predictors is only limited by the imagination of the researcher. However, time does not permit every possible combination of all possible predictors to be tested. The set of predictors to be tested should be limited to those that can be reasonably associated with clouds. Wilks (1995) briefly discusses the problems that can arise from using predictors in a statistical model that cannot be reasonably related to the predictand. Other criterion for choosing predictors includes their availability. For that reason, predictors for this study are limited to those available from the output of the GFS NWP model, or predictors that can be easily calculated from the GFS output. Furthermore, since the SCFM operates on multiple levels of the atmosphere, the predictors should be available on these same levels.

9. Predictor Types

For this study, the predictors are classified by three predictor types. . The purpose of separating predictors by type is to organize them by the type of information they each represent. The first type includes predictors which can be considered as vertical coordinates. These predictors are pressure, geopotential, potential temperature, and temperature. While temperature is not typically used as a vertical coordinate in meteorological contexts, temperature is closely correlated to the vertical structure of the atmosphere, so it is classified as a vertical coordinate here. Relative humidity, cloud water, and RH^4 are considered moisture type predictors. Absolute vorticity, relative vorticity, and vertical velocity are classified as dynamic type predictors. The product of relative humidity and vertical velocity is considered to be both a moisture type and a dynamic type predictor

THIS PAGE INTENTIONALLY LEFT BLANK

III. THE RESEARCH AND OPERATIONAL SCFM

A. INTRODUCTION

In this chapter the working details of the SCFM and the methods used to compare sets of predictors are discussed. First, the theoretical basis for the SCFM model is presented. This will include a description of how a lookup table (LUT) is constructed, and how it is used to diagnose cloud. Then, verification techniques used in this study are shown. Next, discussions of the operational SCFM and research version of the SCFM, the SCFM-R, highlight the specifics of each model and how and why they differ. Finally, the distributions and discretization of the predictors used are given.

B. THEORETICAL BASIS OF THE SCFM

1. Basic Framework of the SCFM

The SCFM operates under a few basic assumptions about cloud cover, NWP models, and the WWMCA. The first is that the formation, persistence, and dissipation of clouds are dependent on sub-grid scale properties of the atmosphere, or properties that are not even modeled. The scale of the grid is determined by the underlying NWP model, in this case the GFS, since this is the scale where the actual dynamics of the atmosphere are modeled. Examples of properties that affect cloud cover which are sub-grid scale, or not modeled at all, include turbulence, entrainment, and the presence of cloud condensation nuclei. As a result, cloud cover cannot be absolutely determined from the bulk meteorological properties that are directly modeled. However, since the bulk state of the atmosphere will also affect cloud cover, a given state of the bulk atmosphere will be associated with a distribution of cloud cover values.

One may try to model the relationship between cloud cover and bulk meteorological properties by using the average of the cloud cover distribution, but this will yield a noisy dataset and a degraded model (McDonald, 2003). It can be shown that using the average of the distribution when diagnosing cloud cover will minimize the RMSE; however, previous modeling efforts have shown that minimizing the RMSE often does not produce realistic cloud forecasts (Norquist et al., 1997; Norquist, 2000). Retaining more information than the average of the cloud cover distribution will allow

higher quality cloud cover diagnoses to be made by accounting for the variability of cloud cover under given bulk meteorological conditions.

Statistics describing the distribution of cloud cover, given a bulk state of the atmosphere, are stored in a LUT. Indices of a cell in the LUT are calculated by binning the bulk meteorological parameters, or predictors, into categories based on their numeric value. Each cell of the LUT contains statistics that describe the distribution of cloud cover associated with the bulk properties of the atmosphere as categorized by the cells indices. Typical statistics to store in a cell include the number of observations corresponding to that cell, the sum of their cloud values, and the sum of the squares of their cloud values. The mean and standard deviation are easily calculated from these quantities and can be used to fit any distribution that can be fit by the method of moments (Wilks, 1995).

The developers of the SCFM at AFWA found that the two-parameter beta distribution fit cloud cover data well (McDonald, 2003). The two-parameter beta distribution is defined as,

$$\beta(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1} \tag{3}$$

where $0 < x < 1$, Γ is the gamma function, and $p, q \geq 0$ are the shape parameters. It is a natural choice for two reasons. First, cloud cover ranges from 0-100%, or 0-1, and the standard form of the two-parameter beta distribution is defined only on the domain (0, 1). Second, the beta distribution can have its peak value anywhere in that domain, and can even be bimodal if the peaks are at zero and one. Having the same domain as the data and being flexible in that domain makes the two-parameter beta distribution a natural choice for the SCFM.

2. Building the SCFM Look Up Tables

Ideally the WWMCA would have full three-dimensional cloud information so that at every model grid point the predictors could be associated with an observed cloud value; however, this is not currently the case. The most reliable information in the WWMCA for the SCFM is the total cloud cover; AFWA conducted experiments using the WWMCA data on four floating levels and the resulting cloud diagnosis was poor

(McDonald, 2005). Better cloud diagnoses are accomplished when the SCFM infers the vertical cloud distribution from the predictors or by another means.

To infer the vertical distribution of cloud, an initial diagnosis of cloud is made at each point in the vertical profile, and then each cloud amount is adjusted to be consistent with the observed total cloud. The initial diagnosis can be made in one of two ways. The first, and preferred, method is to use the standard diagnosis procedure with the LUT as described below. Otherwise, another, perhaps empirical, method is used as an initial condition. In fact, both are used; when the LUT cannot provide a diagnosis due to lack of observations, an empirical method based on equation 2 in Chapter 2 is used. This empirical method was calibrated at AFWA specifically for use in the SCFM, and the critical relative humidities Rh_{00} and Rh_s depend on temperature (Table 1). After the LUT is populated with enough observations, it is used for the initial diagnosis. This means that the empirical method is an initial condition that must be learned and improved upon by the SCFM.

After the initial diagnosis, the cloud amounts must be adjusted to be consistent with the observed total cloud. AFWA originally experimented with several schemes, but settled on the following (McDonald, 2005). For each level, the cloud amount is set equal to the smaller of the total observed cloud amount or the initially diagnosed cloud amount. As the process continues, the LUT will be updated so as to not over diagnose cloud. When updating the LUT, not every point in the vertical is used. The level with the maximum cloud and all levels above it are the only levels used to update the table. In areas of clear skies this includes every level. The reason for this is that the total cloud is often 100%, and the levels below the maximum cloud are not resolvable by the WWMCA, and do not contribute directly to the total observed cloud. As a result, there is no observed information about the cloud amounts below the layer of maximum cloud in general.

Critical Relative Humidity as a Function of Temperature

T	Critical RH (%) Clear	Critical RH (%) Overcast	T	Critical RH (%) Clear	Critical RH (%) Overcast
195	99.482	99.702	260	72.755	85.252
200	99.957	100.000	265	67.080	80.638
205	99.855	100.000	270	67.095	80.263
210	100.000	100.000	275	68.117	80.196
215	100.000	100.000	280	70.700	80.837
220	98.137	99.038	285	73.784	83.677
225	97.381	98.569	290	75.718	85.846
230	97.883	99.335	295	82.892	89.826
235	94.232	96.316	300	60.415	73.892
240	94.716	97.567	305	36.749	48.293
245	89.962	94.164	310	22.262	30.262
250	87.367	94.052	315	12.510	18.915
255	80.181	90.457	320	8.751	9.500

Table 1. Critical relative humidities Rh_{00} and Rh_s used in equation 2 as a function of temperature. These are used as the initial condition or backup method to diagnose cloud when there are not enough observations in the look up table.

3. Diagnosing Cloud with the Look Up Table

Diagnosing cloud from the LUT is a simple task. First the indices of the required cell in the LUT are calculated from the predictors. Then the shape parameters of a two-parameter beta distribution are calculated using the method of moments (Wilks, 1995). A random deviate from this distribution is used as the cloud cover diagnosis for this point. Sometimes this routine will not work. For example, if the mean or standard deviation of cloud values is zero, then the method of moments will fail. This will happen when all of the observed cloud values are the same value; most often when this occurs the cloud amounts are all zero. In these cases, there is no observed uncertainty in the cloud amount associated with the local bulk meteorological parameters, so the mean cloud value for that cell is the diagnosis.

4. Calculating Total Cloud Cover

The only observed quantity available for verification tests is the total cloud cover from the WWMCA. Total cloud cover is used as the "ground truth" for building the LUT, therefore total cloud cover must be a product of the SCFM for verification purposes. However, the SCFM directly diagnoses cloud on many levels of the underlying NWP model, not total cloud cover. This means for verification purposes, total cloud cover must be calculated from the diagnosed cloud field. There are three general

ways that clouds can be stacked in the vertical: perfectly stacked where one cloud lies directly above the other; perfectly not stacked where one cloud is in the spaces left by higher or lower layers, with minimal overlap; and the most general case, random stacking. For the perfectly stacked case, the total cloud is the maximum of the layer clouds, and for the perfectly not stacked case it is the sum of the layer clouds. When randomly-stacked clouds are combined into one layer, the total cloud cover is given by

$$CC = 1 - \prod_i (1 - CC_i) \tag{4}$$

where CC is the total cloud cover, i represents each cloud layer, and CC_i is the layer cloud cover (McDonald, 2005). A cloud layer may vertically extend over several model layers, such as with towering cumulus clouds, and in this case perfect stacking makes sense. However, independent cloud layers will exhibit random stacking. In the SCFM, all cloud layers are treated as independent. This may cause error; however, the expected error is very small since the great majority of total cloud observations from the WWMCA are either 0 or 100%.

C. VERIFICATION OF SCFM CLOUD DIAGNOSIS

Total cloud from the WWMCA is the only observed quantity to verify the SCFM against in this study. First, given the GFS analysis, the SCFM diagnoses cloud and calculates the total cloud. Next, the diagnosed total cloud value is compared to the observed total cloud value with four statistical measures of accuracy. These are the root mean square error (RMSE), bias, correlation of observed total cloud to diagnosed total cloud, and true skill score (TSS) from a contingency table analysis. The RMSE, bias, correlation, and TSS are standard statistics (Wilks, 1995). For the TSS, a contingency table of diagnosed vs. observed cloud cover is constructed with two categories of cloud cover; the first category is cloud cover less than 25%, the second category is cloud cover greater than or equal to 25%. This was done to measure the ability of a predictor set to detect significant cloud cover, where significant cloud cover has been arbitrarily chosen as 25%. This analysis could easily be repeated with a different threshold for significant cloud cover to suit one's operational requirements. Since no single number can adequately measure the fitness of a set of predictors, sets of "good" predictors should be chosen by looking at several such measures of accuracy. The final decision about the

fitness of a set of predictors should include a visual inspection of the cloud diagnoses using those predictors to ensure a realistic diagnosis has been achieved.

Additionally, statistics indicating the validity of the verification were tabulated. During the diagnosis process, errors can occur that must be dealt with; for example a cell in the LUT may not have enough observations to give a good estimate of the cloud distribution. When this happens, there must be a fallback method to diagnose the cloud. If the fallback method is used too often, then the verification will not reflect the true fitness of the predictors. To ensure this is not the case, for each set of predictors the percent of diagnosed points that had an error and used a fallback method is calculated also. These results are not shown here, but were used for a self check during testing. No set of predictors had more than 9% of the points in the verification set that corresponded to a cell in the LUT without enough observations.

D. THE OPERATIONAL SCFM

The first step in running the operational SCFM is preprocessing all of the input data. First, interpolation of GFS fields from a one-degree global grid to two polar stereo graphic 16th mesh (24 km) grids using bicubic interpolation (Hoke, Dr. James E. et al., 1981). Since the SCFM uses a 16th mesh horizontal grid, it is referred to in operation as the SCFM16. AFWA uses two vertical coordinate systems: pressure levels defined in the GFS; and height above ground level. Consequently, interpolation in the vertical is also required for the second grid. The WWMCA is already on a 16th mesh grid, so no interpolation is required; however, some simple quality control is required. One of the fields included in the WWMCA is a time stamp indicating the how old the data at a given point is relative to the valid time of the analysis. This is used to discard any satellite data older than 30 minutes compared to the valid time of the GFS data for the calibration phase.

Building the LUT is done almost exactly as described above, except when diagnosing cloud from the look up table. Instead of using a random deviate from the cloud amount distribution for diagnosing cloud from the LUT, the mean value of the cloud amounts is used. This change is only made for the building of the LUT, cloud diagnoses done for output or verification use the standard algorithm described above. The predictors for the LUT in the operational SCFM are pressure, temperature, pressure

vertical velocity, and relative humidity. They have 9, 26, 11, and 101 categories respectively.

Cloud diagnosis is done as described above for the SCFM in general, except there is a special routine to diagnose cloud if the distribution is too bimodal. If the observed cloud distribution is too bimodal, then a Bernoulli distribution is used instead of a two-parameter beta distribution. Also, if there are not enough observations in a cell of the LUT, the zero cloud cover is diagnosed for that point (McDonald, 2005); this is only done for diagnosing cloud, not when updating the LUT. While this may seem like a very strong assumption, it is expected to have little impact as time goes by since more observations are put into the table, and this error will occur less frequently.

E. THE RESEARCH SCFM (SCFM-R)

1. Set up

The research SCFM (SCFM-R) also starts with interpolation of the GFS fields. However, the research version of the SCFM uses only the pressure levels of the GFS. The interpolation and quality control are done exactly as they are for the operational SCFM, including working on a 16th mesh grid (Hoke, Dr. James E. et al., 1981). Building the LUT is accomplished exactly the same way as it is done for the operational SCFM.

The SCFM-R also addresses some issues that arise with the two parameter beta distribution. The beta-distribution can be so bimodal that zero or one have the same likelihood of being diagnosed, and since the uniform distribution is a special case of the beta-distribution, sometimes it can be very nearly uniform. When either of these situations arises in the model, it may be due to choosing poor predictors, a bad initialization process, or a poor choice of categories for the predictors. This condition results in a wide range of uncertainty, which leads to some points being diagnosed as having too much cloud cover. If only one or two of these bad points are in a column, it will over-diagnose total cloud cover if it is not already 100%. When this error is observed it is often so bad that it causes a diagnosis of 100% cloud cover over the entire globe. To prevent this, AFWA simply tuned the number of categories of each predictor in the operational SCFM to minimize this effect (McDonald, 2005). However, when

comparing different sets of predictors, the resources and time are not practically available to tune every possible set of predictors.

In order to filter out this noise, bad cloud diagnoses had to be detected and handled separately. Detection is possible by noting the values of p and q, the parameters of the two-parameter beta distribution, which correspond to distributions that are too uniform (Fig. 1A) or too bimodal (Fig. 1B). Under the special condition that p = q = 1 (Fig. 1), the two-parameter beta distribution is the uniform distribution, and if p, q < 1 the distribution is bimodal. Also, the regions of the p-q space that correspond to nearly uniform and too bimodal overlap significantly (Fig. 1C).

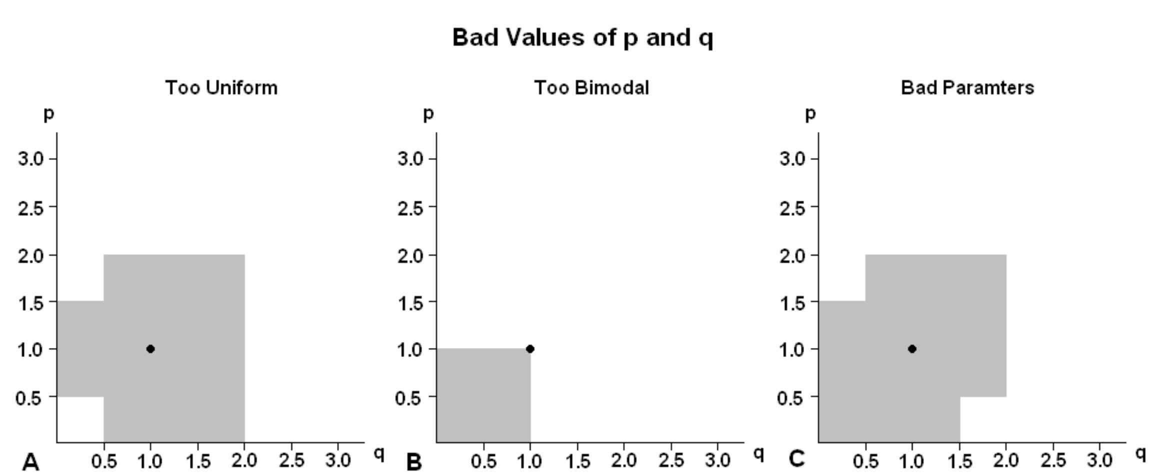

Figure 1. Values of p and q that correspond to uncertain cloud diagnoses. A) Parameters that correspond to a distribution that is too nearly uniform. B) Parameters that correspond to a distribution that is too bimodal. C) The full set of bad shape parameters. The black dot at p=q=1 corresponds to the special case where the two parameter beta distribution is the uniform distribution.

Once a bad diagnosis has been detected by checking p and q, it is handled in one of two ways. First, it can be ignored, this will be called diagnosis scheme 1 (DS1), which is exactly how the operational SCFM handles it. The second way to handle it is to fall back to the initial cloud diagnosis method (DS2). Either way, as time goes on and more observations are recorded in the LUT, the fall back method will be used less and less, eventually having a negligible impact. Results will be presented for both cases.

Another minor difference arises when combining cloud layers. The operational SCFM only diagnoses cloud on 9 levels, whereas the research SCFM diagnoses cloud on 19 levels. The reason for this difference is that AFWA only has 9 categories for pressure, and since the model is run on pressure levels, this limits the number of levels that can be used. Furthermore, the lowest layer of the final output of the operational SCFM is taken from the version that uses height as a vertical coordinate and includes most of the boundary layer. To be consistent with the operational SCFM, instead of combining 19 layers of cloud, the research version combines 9 layers for verification. Using the mean layer cloud between 1000 mb and 975 mb for the boundary layer, the 900 mb, 800 mb, 750 mb, 650 mb, 550 mb, 450 mb, 350 mb, 250 mb, and boundary layer clouds are combined to make the total cloud amount. The 900 mb layer and above are the same layers used in the operational SCFM.

2. Predictors and Data

Possible predictors were limited by the output of the GFS, and resources limited the number of predictors that could be tested. The GFS data used in this study were the final analysis available from the NCEP public FTP site, and it has eight fields that are on nineteen levels. These fields are: geopotential, temperature, pressure vertical velocity, relative humidity, absolute vorticity, zonal wind speed, meridional wind speed, and cloud water. Since the GFS data is on pressure levels, pressure is implicitly included in the data and is thus included as a predictor. Potential temperature, wind speed, relative vorticity, and various transformations and combinations of all the aforementioned predictors can be calculated and used as predictors too. Relative humidity to the fourth power and the product of relative humidity and vertical velocity are examples of such predictors (Norquist et al., 1994; Norquist et al., 1997; Norquist, 1999; Norquist, 2000; Weiss, 2001). Thus far, 14 predictors have been explicitly listed, and choosing sets of predictors that perform well in the SCFM requires that all possible combinations are compared. Therefore 1001 different LUT's must be built and verified; however, enough computer resources were not available to complete this in a reasonable time. As a result, zonal wind speed, meridional wind speed, and wind speed were not used as predictors, reducing the number of LUT's required to 330. For this study, dynamically it could not be

justified to include wind predictors given the computational cost. A more detailed discussion of the other predictors is given in Chapter II.

The actual data used for this study covered the one week period of August 20 to August 26, 2005. A second period was used to ensure that there was enough data to populate the LUT, so the final results reported in this thesis also used data from December 1 to December 14, 2005. Since the data has global coverage, both the summer and winter seasons are represented. Fall and Spring are not represented in the data, though AFWA has not reported any change in performance in different seasons.

3. Distribution of Predictors and Predictor Categories

Choosing how to discretize the predictors into numbered categories is very important. If the categories are too large, then atmospheric conditions corresponding to cloudy skies will be placed into the same category as conditions corresponding to clear skies. This causes unnecessary uncertainty for the cloud amount associated with a given LUT cell, and thus a generally broader distribution of cloudiness values will be in the cell, ultimately leading to a poorer cloud diagnosis. Conversely if the categories are too small, then it will take more training data to populate the LUT, and thus it will take longer to build an LUT that provides a skilled cloud diagnoses. One must also consider memory and storage requirements for the computer system running the model. Since there are four predictors for each LUT, if each predictor has N categories, the number of cells in the LUT will be N^4, so doubling N will require 16 times more storage and memory for a LUT. For this study, categories ranged from 0 to 50 for each predictor. Six days of data, each with four GFS analyses, were used to build the LUT, and it was verified with the seventh day's data. Later, to ensure enough observations were in the tables, another 13 days of were added to the LUT's, and verification was done on the fourteenth day.

Upper and lower bounds were also necessary for the predictors. Any value of a given predictor below the lower bound falls into the lowest category, and any value above the upper bound falls into the highest category. The bounds for this study (Table 2) were chosen subjectively by inspecting the distribution of each predictor (Fig. 2), and placing the bounds so that any portion of the distribution with a significant relative

frequency was within those bounds. In Figure 2, the large dots on the horizontal axis show the upper and lower bounds for the categories.

Some of the distributions in Figure 2 are multimodal due to the vertical coordinate system of the output of the GFS. Output from the GFS used for this study was on pressure levels, and all of the predictors were available on 19 specific levels. Hence, there are 19 peaks on the histogram for pressure, each corresponding to a pressure level. This means that many of the cells of the LUT had no observations, since there were no pressure values corresponding to their category. Operationally, the 50 categories of pressure would be reduced to 19 or fewer to save space. Geopotential and temperature are also multimodal (Fig. 2), presumably because of their strong correlation to pressure. The Spearman rank correlation matrix for the predictors (Table 3) shows the correlation of the predictors. The large correlations of pressure, geopotential, potential temperature, and temperature to each other validate this presumption, as well as justify them being treated together all as vertical coordinate type predictors.

Bounds for Categorizing Predictors

Predictor	Lower Bound	Upper Bound
Pressure	200 mb	1,000 mb
Geopotential	0 gpm	10,000 gpm
Temperature	215 K	315 K
Vertical Velocity	-0.5 mb/s	0.5 mb/s
Relative Humidity	0.0 %	100.0 %
Absolute Vorticity	$-2.0\text{x}10^{-4}$ /s	$2.0\text{x}10^{-4}$ /s
Cloud Water	0.0 kg/kg	$1.0\text{x}10^{-4}$ kg/kg
Potential Temperature	280 K	360 K
Relative Vorticity	$-3.0\text{x}10^{-4}$/s	$3.0\text{x}10^{-4}$/s
(Relative Humidity)^4	0.0 %	100.0 %
Relative Humidity x Vertical Velocity	-25.0 % mb/s	25.0 % mb/s

Table 2. Upper and lower bounds used when assigning categories to predictors. The corresponding points in the predictor distributions are shown in Figure 2.

Spearman Rank Correlation Matrix for SCFM Predictors

	P	Φ	T	ω	RH	ςA	Cloud Water	θ	ς	RH^4	RH x ω
P	1.000										
Φ	-0.997	1.000									
T	0.831	-0.799	1.000								
ω	0.041	-0.038	0.000	1.000							
RH	0.417	-0.436	0.170	-0.214	1.000						
ςA	0.005	0.026	0.232	-0.079	-0.046	1.000					
Cloud Water	0.085	-0.094	-0.008	-0.209	0.549	-0.053	1.000				
θ	-0.783	0.818	-0.380	-0.066	-0.563	0.245	-0.210	1.000			
ς	0.018	-0.011	0.019	-0.009	0.038	0.298	0.060	0.000	1.000		
RH^4	0.417	-0.436	0.170	-0.214	1.000	-0.046	0.549	-0.563	0.038	1.000	
RH x ω	0.072	-0.071	0.005	0.947	-0.133	-0.078	-0.197	-0.106	-0.006	-0.133	1.000

Table 3. Spearman rank correlation matrix for the SCFM predictors.

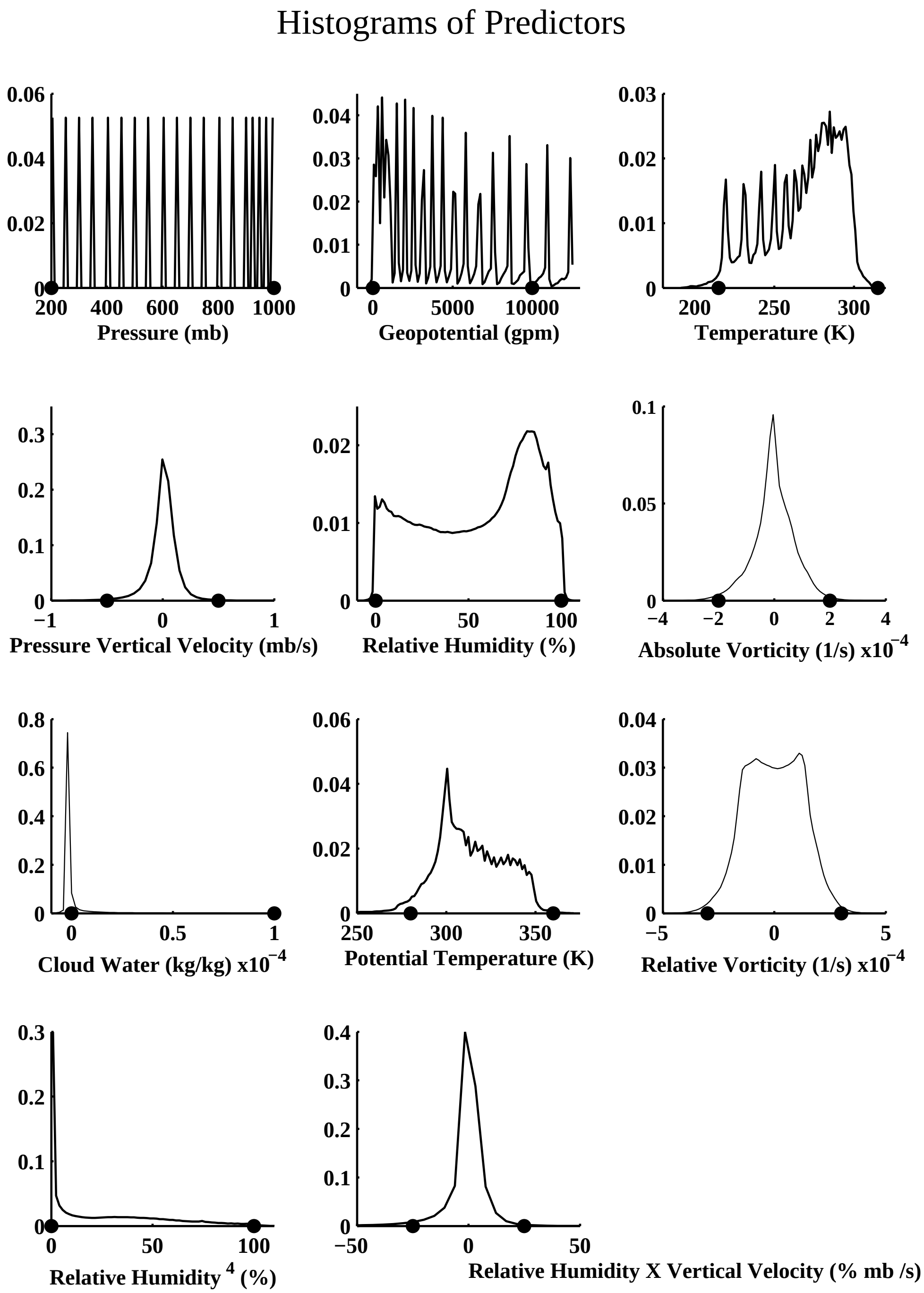

Figure 2. Histograms of the predictors for 0000Z, August 20, 2005. The dots are the upper and lower bounds for the predictor categories given in Table 2.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

A. OVERVIEW

The results of the experiments with the SCFM-R are now presented. First, the top performing sets of predictors are described, as are their characteristic errors. This includes a comparison of DS1 and DS2. Following is a comparison of verification statistics for DS1 and DS2, a comparison between the SCFM-R and the operational SCFM as reported by AFWA (McDonald, 2005; Cloud Models Team, HQ AFWA/DNXM, 2005). Finally, geostationary satellite images are compared to the total cloud cover from the WWMCA and some total cloud cover diagnoses from the SCFM-R with DS2.

B. VERIFICATION OF PREDICTOR SETS

Top performing sets of predictors were chosen by ranking them three ways, then taking the average of those three ranks. While this is not an established procedure, it was effective for placing sets of predictors which performed well under all three ranking methods at the top of the list. The three ranking methods were by: lowest RMSE, highest correlation, and highest TSS. It is not claimed in this thesis that the absolute order of the ranking determines an absolute difference in fitness for the sets of predictors; instead it only places “good” predictor sets near the top of the list. For this reason, results given in tables 3 and 4 are ordered by their RMSE. Also, only the top 15 sets from each diagnosis scheme are shown.

1. Verification of DS1

Top performing sets had predictors of all three types for DS1 (Table 4). Pressure was the most common vertical coordinate, followed by geopotential height. Potential temperature was only among the top 15 sets of predictors in one set, and temperature was completely absent. Every top performing set had either relative humidity, or relative humidity to the fourth power, highlighting the importance of relative humidity to the SCFM and clouds. Additionally, five sets also had cloud water, highlighting the importance of moisture to clouds in the SCFM. For the dynamic predictors, vorticity was most frequently among the top performing sets, but vertical velocity was also common. Most sets either had two moisture predictors, or two dynamic predictors. The product of

vertical velocity and relative humidity was considered both a dynamic and a moisture type predictor.

Top Performing Predictors (DS1)

Vertical Coordinates				Moisture				Dynamic			DS1			
P	Φ	T	θ	RH	Cloud Water	RH^4	RHxω	ςA	ς	ω	RMSE	R	Total Skill Score	Bias
√						√	√		√		39.28	0.48	0.30	-4.9
√						√			√	√	39.47	0.49	0.32	-6.9
√						√	√	√			39.69	0.47	0.30	-4.9
√						√		√		√	39.96	0.47	0.32	-6.8
√						√		√	√		40.05	0.46	0.31	-6.3
	√					√		√		√	40.31	0.46	0.30	-6.8
√				√			√		√		40.46	0.46	0.31	-9.0
√				√					√	√	40.93	0.47	0.34	-11.8
√				√				√		√	41.67	0.44	0.34	-11.6
	√		√			√				√	41.73	0.44	0.35	-11.5
√					√	√			√		42.13	0.46	0.40	-14.7
√					√	√		√			42.57	0.44	0.38	-14.4
	√				√	√			√		42.65	0.45	0.37	-15.3
√					√	√	√				42.72	0.46	0.41	-16.0
√					√	√				√	42.74	0.46	0.42	-16.6
√		√		√						√	43.84	0.39	0.30	-15.3
Initial Conditions											49.92	0.41	0.23	24.3

Table 4. Top 15 performing predictor sets using Diagnosis Scheme 1 (DS1).

Top Performing Predictors (DS2)

Vertical Coordinates				Moisture				Dynamic			DS2			
P	Φ	T	θ	RH	Cloud Water	RH^4	RHxω	ςA	ς	ω	RMSE	R	Total Skill Score	Bias
√						√			√	√	40.78	0.51	0.39	-2.3
√						√		√		√	40.91	0.51	0.40	-1.6
√						√		√	√		41.04	0.50	0.39	-2.4
√						√	√		√		41.26	0.51	0.38	3.3
	√					√			√	√	41.32	0.51	0.38	3.9
√						√	√	√			41.38	0.51	0.39	2.9
	√					√		√		√	41.50	0.51	0.39	3.1
	√			√			√	√			41.55	0.40	0.39	-1.7
	√					√		√	√		41.68	0.50	0.38	3.2
	√				√	√		√			41.96	0.49	0.40	-3.1
		√	√		√	√					42.19	0.48	0.42	-7.4
	√				√	√			√		42.20	0.50	0.49	-4.0
	√				√	√	√				42.25	0.50	0.41	-4.2
	√				√	√				√	42.39	0.50	0.41	-6.2
√					√	√			√		42.50	0.49	0.42	-7.3
√		√		√						√	44.09	0.42	0.36	-15.4
Initial Conditions											49.92	0.41	0.23	24.3

Table 5. Top 15 performing predictor sets using Diagnosis Scheme 2 (DS2).

Considering the natural dependence of cloud cover on moisture, one may ask if moisture alone is good enough to diagnose cloud cover. Since there were not four predictors that only included moisture, this question cannot be answered directly with this research. However, many sets of predictors had three moisture predictors, RH, RH^4, and cloud water. If moisture alone is a good predictor of cloud cover, then even the worst of these sets of predictors should perform well. This was not observed. Predictor sets with all three moisture predictors had RMSE's ranging from 43.1% to 51.8%. Again, the importance of having more than one type of predictor is apparent, implicit in the dynamical character of clouds.

There is really no single set of predictors, or group of sets, that stands apart from the rest. When the sets are sorted by RMSE in increasing order, the RMSE values smoothly increase from the top performers down to the worst (Fig. 3). Overall, errors ranged from 39-43% for the RMSE, 0.44-0.49 for the correlation, and 0.30-0.42 for the TSS among the top performers. The largest RMSE's overall were on the order of 54%. When the predictors were P, T, ω, and RH, which are the predictors AFWA currently employs, the errors were: on the order of 44% for the RMSE, on the order of 0.40 for the correlation, and 0.30 for the TSS. It is clear that the top performing sets (Table 4) show improvement over P, T, ω, and RH.

2. Verification of DS2

DS2 verified (Table 5) very similar to DS1 (Table 4). The main difference between DS1 and DS2 was the lower magnitude of the bias. DS2 had roughly half the bias of DS1 for the top performers. Comparing other statistics, the RMSE for DS2 was generally slightly higher for the top performers. However, the correlation and TSS were also slightly higher in general. Other than the bias, there is no exceptional difference in the verification statistics between DS1 and DS2. Even the differences in the bias only apply for the top performers. Generally there are no major differences between DS1 and DS2 that apply to all sets of predictors. The top performers for DS2 (Table 5) were similar to DS1 except for a few minor differences. First, there were more sets with geopotential as the vertical coordinate predictor type, and RH^4 was more frequently among the top performers at the expense of RH.

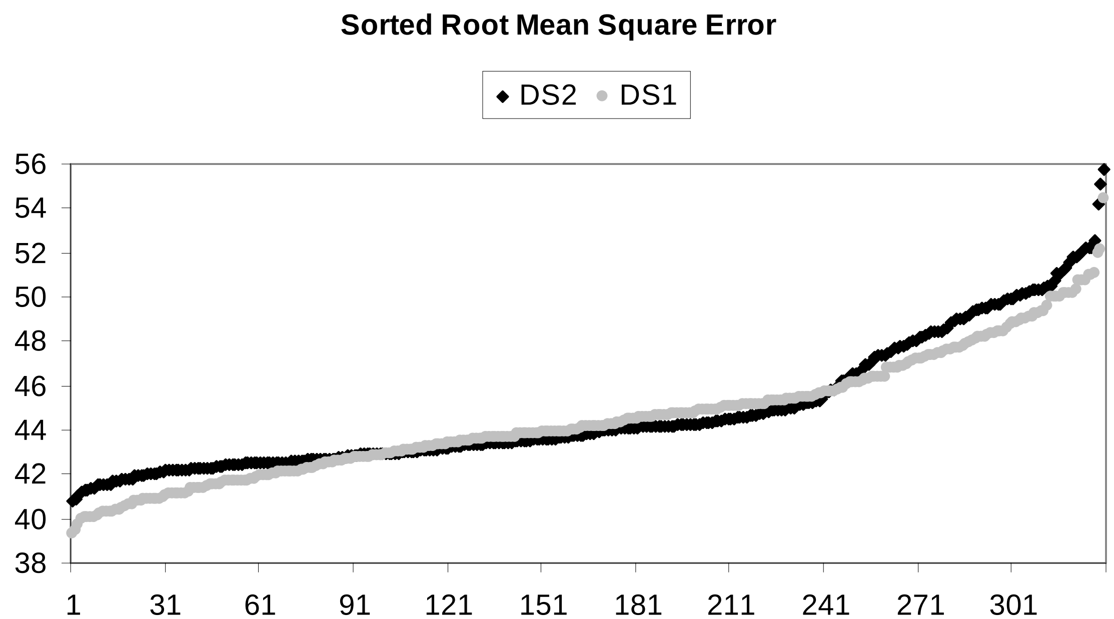

Figure 3. Root Mean Square Error, sorted in increasing order for each of the 330 possible predictor combinations.

C. COMPARISON BETWEEN THE SCFM-R AND THE SCFM

In Chapter III, both the SCFM-R and the operational SCFM were described and their differences highlighted. The impacts of those differences will now be explored. The two main differences were the number of categories for each predictor, and the method of diagnosing cloud cover in the boundary layer. The SCFM-R was run with both DS1 and DS2, and with the same predictors and categories as the operational SCFM. When comparing the model runs where the only difference is the number of categories (Table 6) there is little difference in their performance. Thus, it is concluded that the choice of categories for the SCFM-R had little impact on measuring the performance of sets of predictors.

Comparing DS1 with AFWA categories (Table 6) to the results reported by AFWA for the operational SCFM (McDonald, 2005; Cloud Models Team, HQ AFWA/DNXM, 2005) leaves the boundary layer as the only major difference between the models. From this comparison, it is apparent that the models produce similar cloud diagnoses, at least when comparing overall errors, and that they all outperform the initial condition and a random guess cloud cover diagnosis. However, the operational version

of the SCFM does produce a slightly better RMSE, a much better correlation, and a much better bias (Table 6). Since the major difference is the treatment of the boundary layer, it appears that the operational SCFM has the better method for handling the boundary layer.

Comparison of Diagnosis Schemes

Diagnosis Method	RMSE	R	Bias	Total Skill Score
DS1 AFWA Categories	44.90	0.38	-18.80	0.27
DS2 AFWA Categories	46.36	0.41	-22.93	0.32
DS1	43.84	0.39	-15.28	0.30
DS2	44.09	0.42	-15.38	0.36
Random	66.56	-0.0008	0.05	-0.0011
Initial Conditions	49.42	0.41	24.26	0.23
Operational SCFM	42.50	0.51	~3	--

Table 6. Comparison of diagnosis schemes and number of categories. AFWA results are as reported by AFWA for the operational SCFM (McDonald, 2005).

D. VISUAL COMPARISON OF CLOUD DIAGNOSES

Visual comparisons between satellite imagery, the WWMCA total cloud cover, the initial diagnoses scheme, and a few total cloud cover diagnoses from the SCFM-R will now be discussed. Some important features that will be discussed in this section include the mature system in central Canada (Fig. 4), and tropical convection over equatorial Africa (Fig. 5). Both are clearly visible and well represented in the WWMCA (Fig. 6).

1. Comparison of WWMCA and Geostationary Satellite

Some features of the WWMCA (Fig. 6) are very unrealistic. For example, sharp gradients with a very smooth curved edge appear in the South Atlantic and South Indian oceans. This is attributed to the multi-satellite input to the WWMCA. The long curved edges correspond to the edge of the footprint for geostationary satellites. Outside the footprint, the data from polar orbiting satellites is older, and the cloud cover has changed since the most recent pass. This will not interfere with the SCFM or the SCFM-R, since both keep track of the valid time of each point in the WWMCA and discard any point not valid within half an hour to the GFS analysis.

2. Comparison of Diagnosed Cloud Cover Images

If cloud cover were diagnosed only with the initial condition, the result would be as in Figure 7. The bias for over diagnosing cloud cover, which was apparent in Tables 4 and 5, is also very apparent here. However, comparing this image with Figures 4, 5, and

6, it is also apparent that when the initial condition diagnoses clear skies, it is correct most of the time.

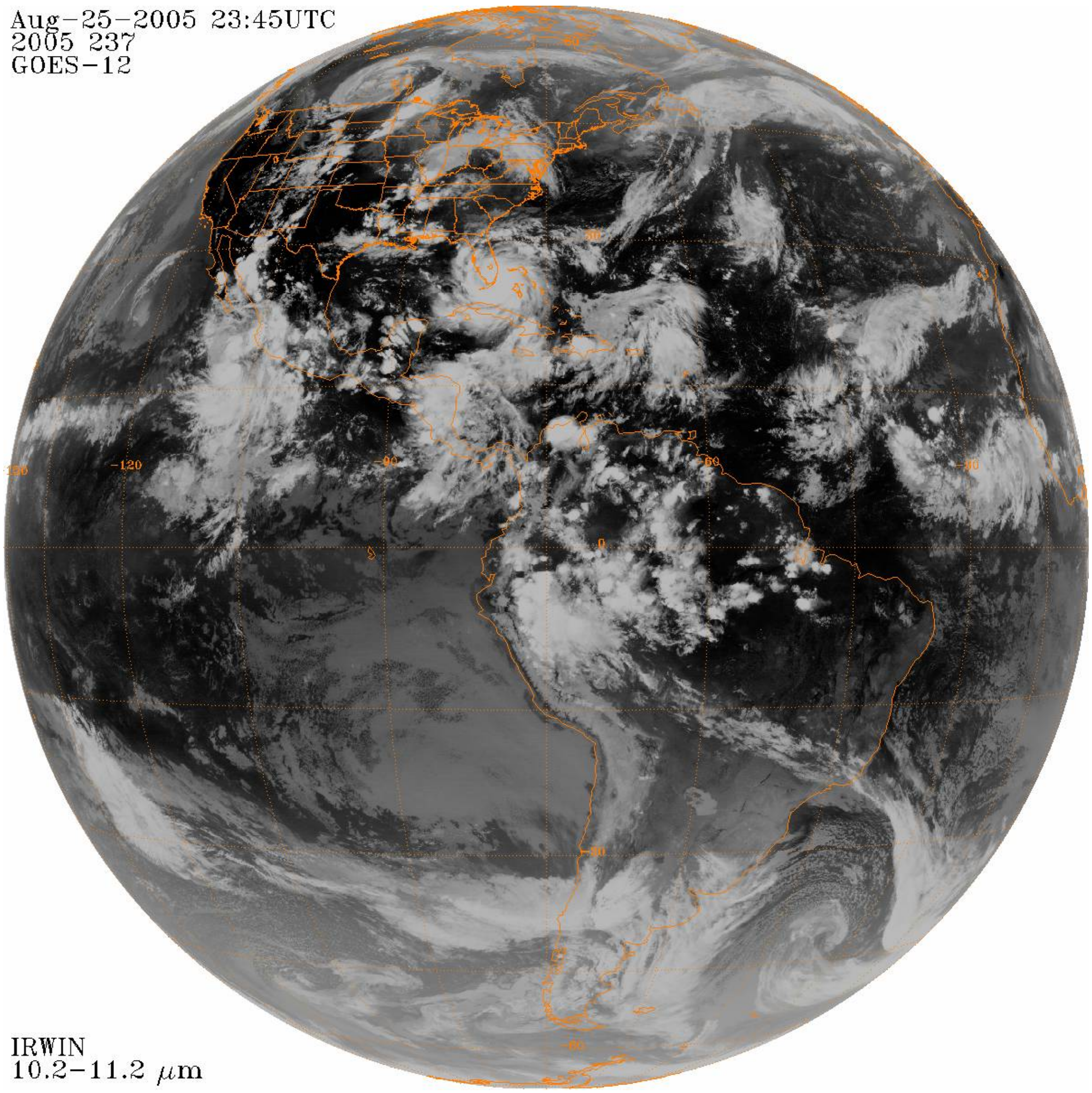

Figure 4. GOES-12 full disk image from 23:45Z, August 25, 2005 (GIBBS: Global ISCCP B1 Browse System. 2006).

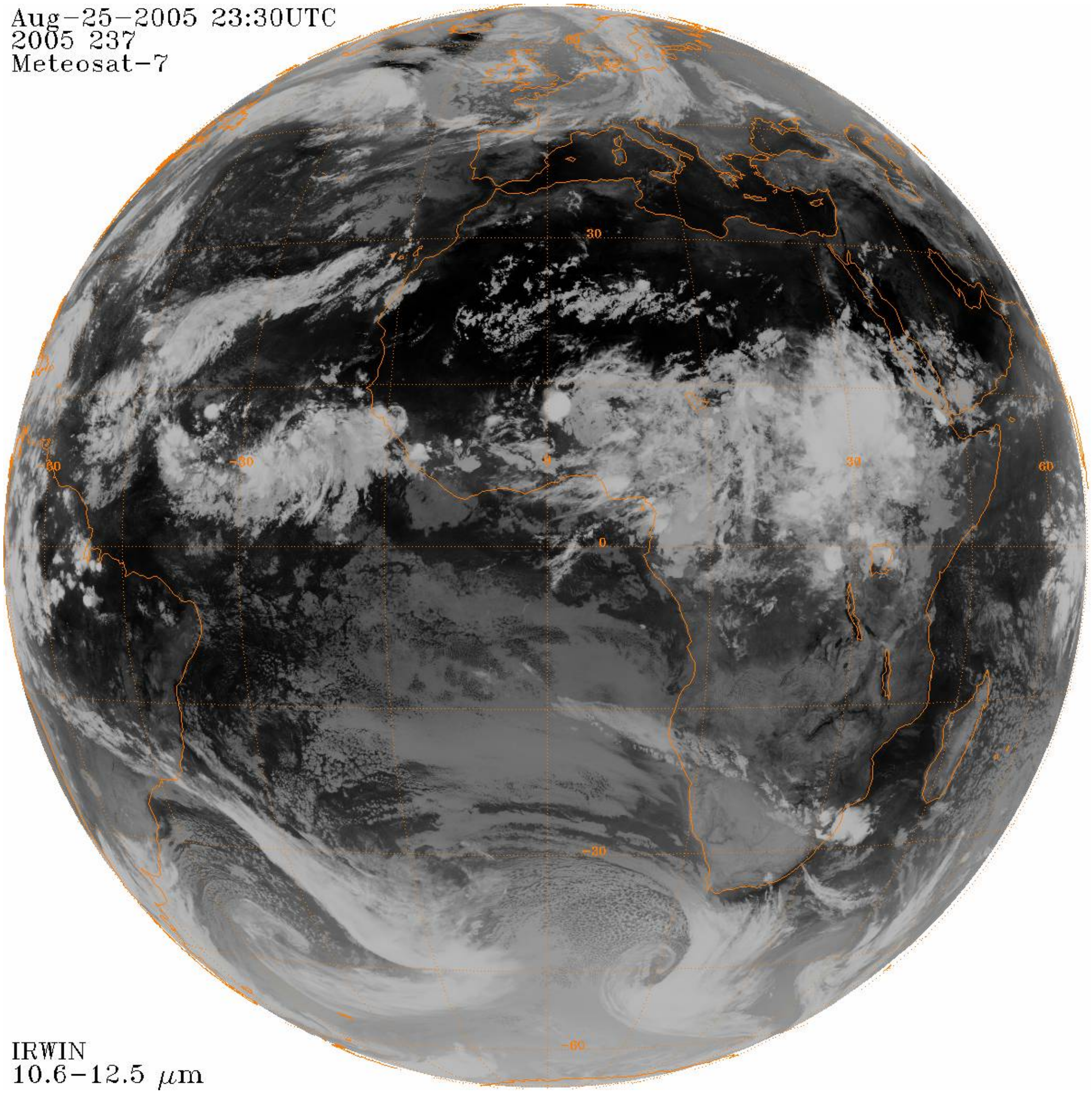

Figure 5. Meteosat – 7 full disk image from 23:30Z, August 25, 2005 (GIBBS: Global ISCCP B1 Browse System. 2006).

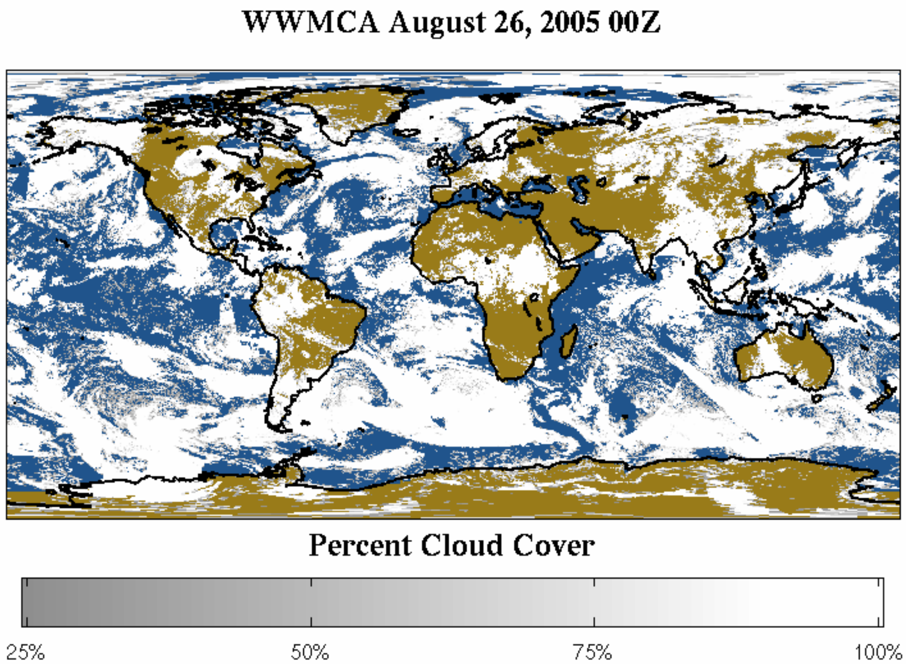

Figure 6. World Wide Merged Cloud Analysis for 0000Z, August 26, 2005.

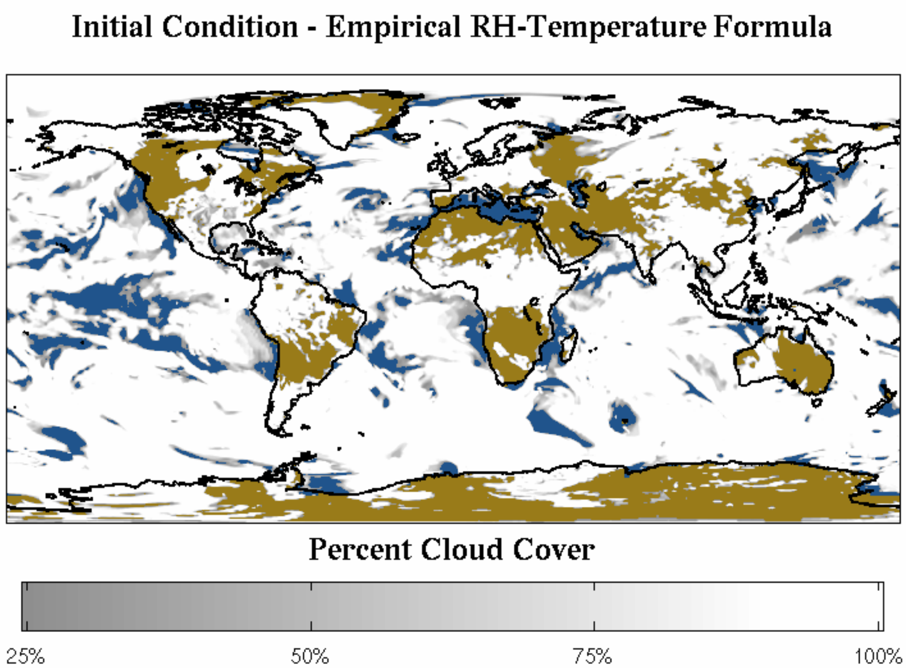

Figure 7. Cloud Cover diagnosed with initial condition from AFWA. The initial condition is only a function of relative humidity and temperature.

When the predictors are pressure, temperature, vertical velocity, and relative humidity (Fig. 8), there is a regional bias. Tropical regions have cloud cover under-diagnosed drastically. Most of the tropical regions have been diagnosed with 25-70% cloud cover, when really there is 100% cloud cover (Fig. 6); note the low values of cloud cover over equatorial Africa. However, in the mid-latitude and polar regions, most of the cloud cover is diagnosed as close to 0% or 100%, which is more accurate when compared to the WWMCA (Fig. 6). This suggests that the large negative bias of these predictors in the SCFM-R (Table 6) is due primarily to the under-diagnosis of cloud cover in the tropics.

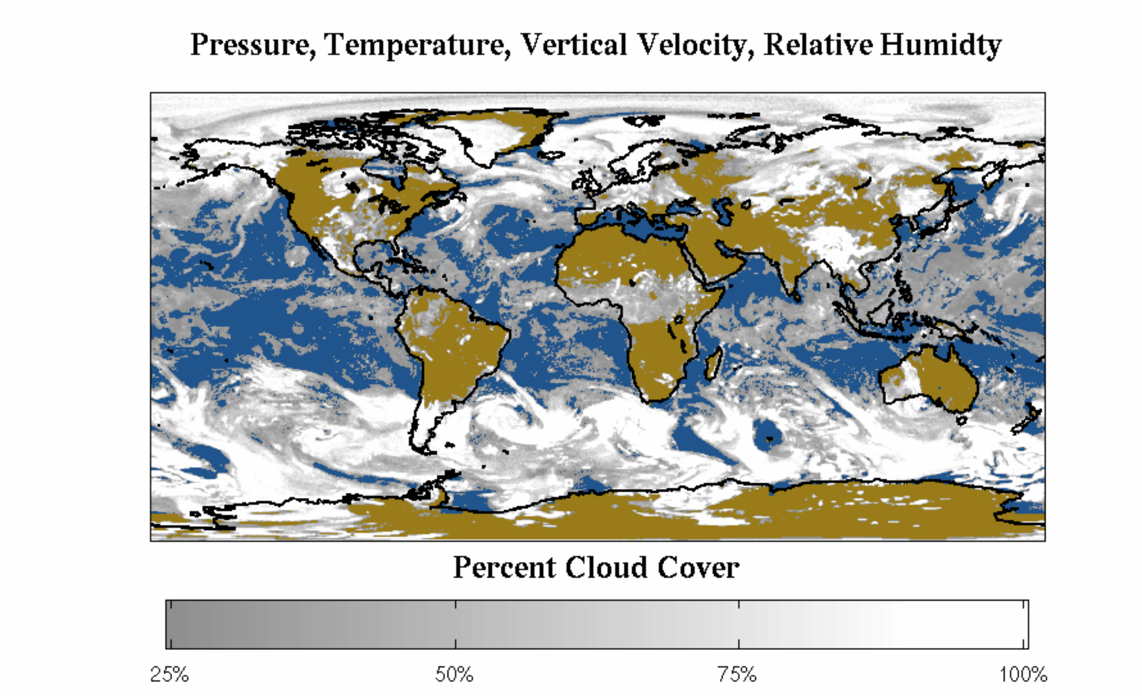

Figure 8. Cloud cover diagnosed from pressure, temperature, vertical velocity, and relative humidity in SCFM-R with DS2.

A different kind of artifact occurs when vorticity is used as a predictor. When absolute or relative vorticity is used as a predictor (Fig. 9), zonal cloud bands are diagnosed. It is not considered coincidence that the bands are parallel to latitude lines. Implicitly, latitude is included in absolute vorticity as the planetary vorticity component. However, this is not included in the relative vorticity predictor. Currently the source of this artifact is unknown. Other than this glaring inaccuracy, vorticity seems to be a very

skillful predictor when considering the verification statistics and the top performing sets of predictors. This should motivate future research to understand the source of this error, so that it can be corrected and the otherwise quality diagnosis of predictor sets with vorticity can be utilized.

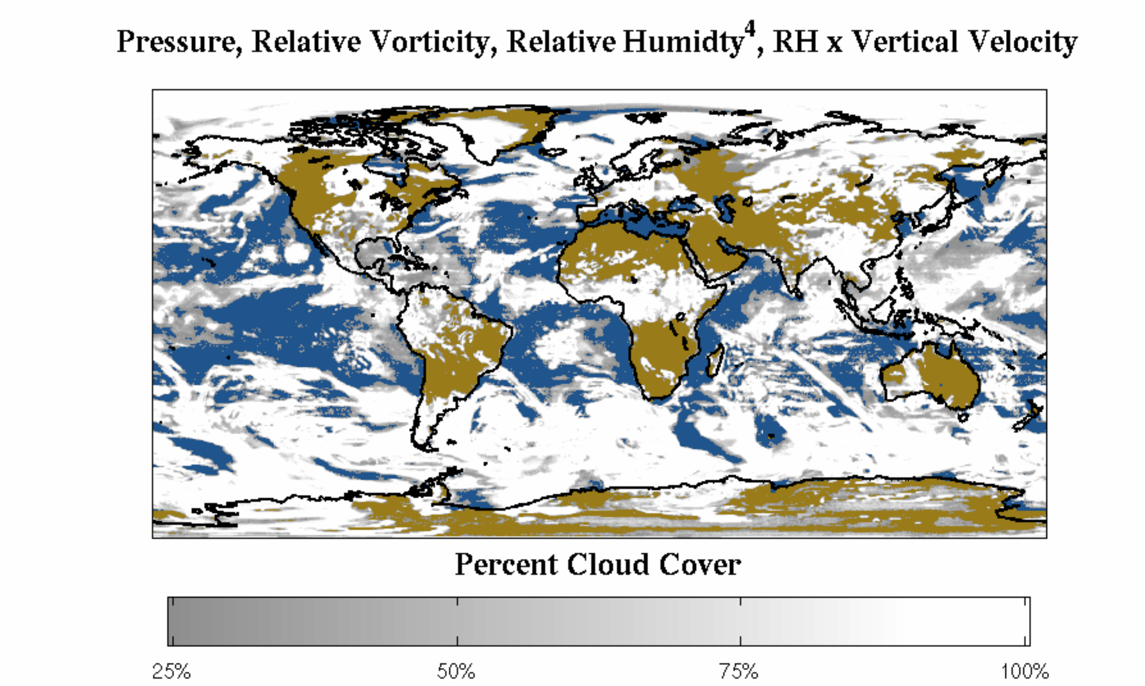

Figure 9. Cloud cover diagnosed from pressure, relative vorticity, relative humidity^4, and the product of relative humidity and vertical velocity in the SCFM-R with DS2.

A typical total cloud cover diagnosis from a top performing set without vorticity is shown in Figure 10. There is no apparent regional bias; notice both the mature system over central Canada and the convection of equatorial Africa are both diagnosed as 100% cloudy. Also, no unrealistic artifacts are present. This predictor set clearly demonstrates skill both with verification statistics (Table 5) and by producing a realistic looking cloud diagnosis.

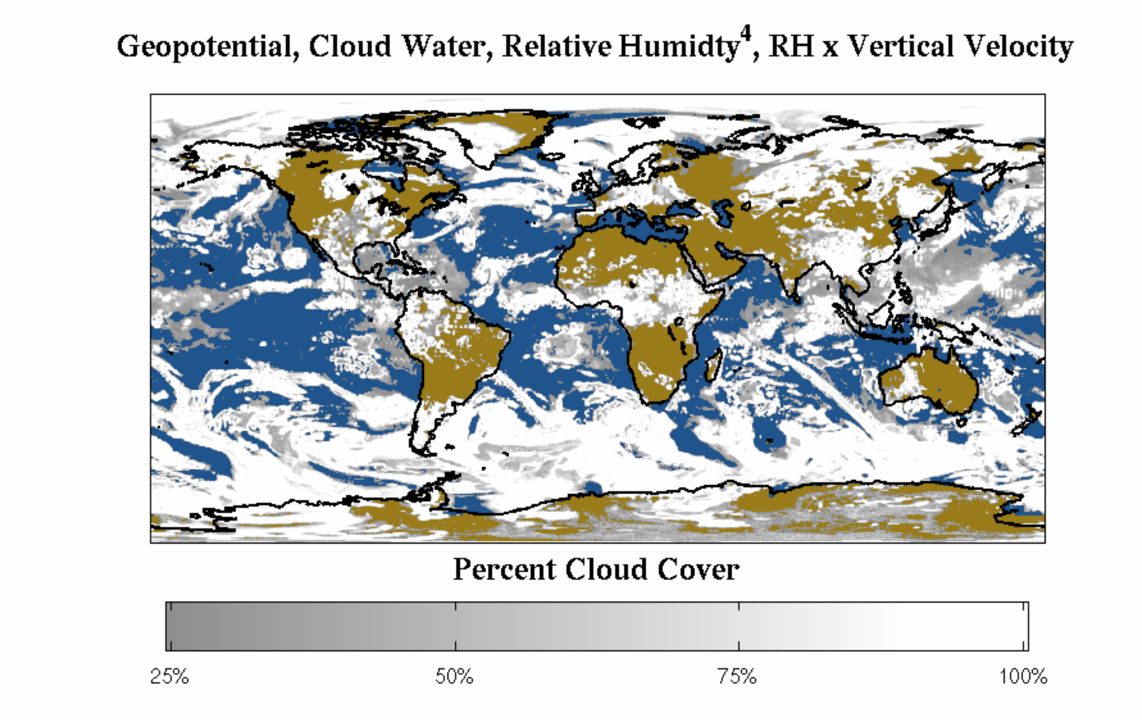

Figure 10. Cloud cover diagnosed from geopotential height, cloud water, relative humidity[4], and the product of releative humidity and vertical velocity in the SCFM-R with DS2.

THIS PAGE INTENTIONALLY LEFT BLANK

V. DISCUSSION

A. CONSISTENCY OF TOP PREDICTORS WITH OTHER MODELS

1. Model Output Statistics

Although the underlying assumptions of the SCFM are significantly different from those of the DCFM and MOS, the top performing predictors from each model are now compared. Since MOS only diagnoses the lowest layers of clouds, diagnosing the base and amount, there are no vertical coordinate type predictors. Recent observations are also used as predictors, however these do not apply to the SCFM, and so only a few of the model data predictors will be listed. For region 20, the inland Southeast US, for the months of October to March, a few of the chosen predictors are: humidity at various levels, vertical velocity, the product of relative humidity and vertical velocity, and wind components (Weiss, 2001). So dynamic and moisture model predictors are present, but no vertical coordinate. However, since the predictors are taken from specific levels, such as the 850 mb relative humidity, there is implicitly a vertical coordinate as a predictor also. Vorticity was available as a predictor for the development of the cloud cover MOS equations, however, it was not selected. It should be noted that MOS includes previous observations to include a persistence component to the diagnosis as well as climatology predictors, so the predictor sets are not directly comparable (Weiss, 2001).

2. Diagnostic Cloud Forecast Model

The DCFM also relies on linear relationships between the predictors and predictand. The vertical coordinate type of predictors is included implicitly, since the model is divided into three layers: low level clouds, middle clouds, and high clouds (Norquist, 2000). Different predictor equations are developed for each layer every time the model is run, so different predictors can be chosen each time. However, Norquist et al (2000) did list some predictors that were chosen most frequently during their latest study. Among these predictors were: zonal or meridional wind speed, relative humidity squared, wind shear, precipitable water, cloud water, maximum of relative humidity in a layer, relative humidity, lapse rate of potential temperature, temperature, sine of longitude, and cosine of latitude (Norquist, 2000). Again, moisture is very important, and

dynamically, wind speed seems to be the only frequently chosen predictor, even though vorticity was available as a predictor.

3. Summary of Comparison to Other Models

From both the DCFM and MOS the importance of moisture as a predictor is apparent as expected. However, wind speed, which was not included in the SCFM for practical purposes, also was a commonly chosen predictor. This suggests that wind speed should have been retained as a possible predictor in the SCFM-R. Conversely, vorticity, which was frequently among the top performers in the SCFM-R, was not selected as a predictor in MOS or the DCFM. However, it is important to remember that the SCFM does not rely on assumptions of linearity, so some differences are to be expected.

B. INITIAL CONDITIONS

One of the largest sources of error for the SCFM may in fact be the initial condition, which is represented by Equation 2 with values of Rh_{00} and Rh_s listed in Table 1. The values in Table 1 were calculated by matching the cumulative probability distributions of relative humidity and cloud cover for each temperature (McDonald, 2005). However, the relative humidity is calculated with respect to liquid water, not ice, and ice deposition occurs at much lower specific humidities than does liquid condensation (Rogers and Yau, 1989). The reason for this is the difference in saturation vapor pressures over ice and water. As a result, when considering relative humidity for liquid water at very cold temperatures, deposition and cloud formation will occur at much lower relative humidities than in warmer temperatures (Rogers and Yau, 1989). So the critical relative humidities Rh_{00} and Rh_s in Table 1 should be much lower at low temperatures. Figure 11 shows a plot of the critical relative humidities vs. temperature; notice the local minima near 270 K. As temperature decreases, so should these values, but in fact they increase to very near 100%. Furthermore, the critical relative humidities at warm temperatures are much too low, since clouds formation above 260 K is typically dominated by liquid cloud droplets (Rogers and Yau, 1989).

Since in the real world clouds will actually form at much lower relative humidities in cold regions, the humidity in regions with these low temperatures will likely never get high enough to initiate cloud cover to be diagnosed. Also, since clouds in warm regions typically form at much higher relative humidities than suggested by

Table 1 and Figure 11, warm regions will likely be over diagnosed with clouds. Taking the thermal structure of the atmosphere into account, high clouds will be drastically under diagnosed and low clouds will be over diagnosed. Further, since cloud cover is partitioned in the vertical initially using the initial condition while building the LUT, the SCFM is being trained with an unphysical and unrealistic vertical cloud distribution.

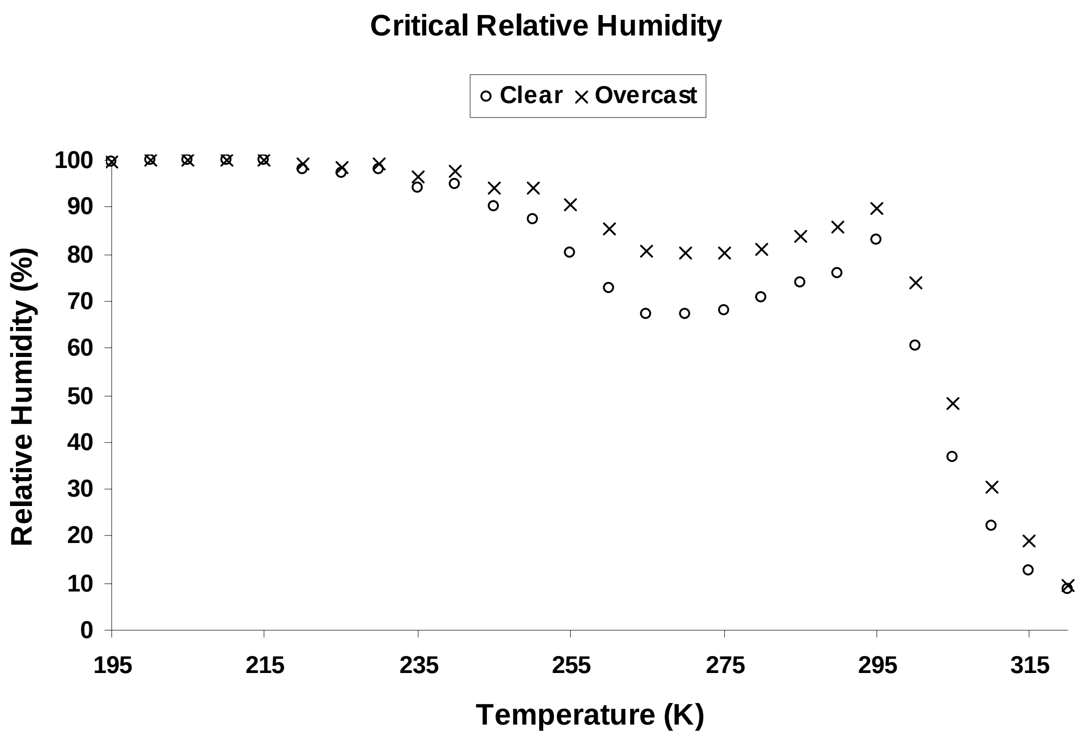

Figure 11. Critical relative humidities for overcast (Rh_s) and clear (Rh_{00}) skies used in equation 2 as the initial condition in both the operational SCFM and SCFM-R.

C. OTHER POSSIBLE PREDICTORS AND LOOK UP TABLE CONFIGURATIONS

One may want to include latitude, longitude, or date into the SCFM as predictors, or even to add more than four predictors to the SCFM, however, this would be very difficult to implement. While adding these predictors may add valuable information, such as season or location dependent effects, it would be impractical to train and use the model. First, since the SCFM is a global model, multiple seasons are always present in the model domain, so including date or season alone would not really add much information. Unless latitude was included as a predictor to separate out each hemisphere, so seasons could really be separated, date would not be a good predictor. Further, since

each season or date would be treated independently, at least a year of data would be needed for training the SCFM. Also, it would take much longer for the SCFM to improve on itself, since at any given time of year it would only have access to part of the LUT. Including latitude and longitude may be useful by allowing the LUT to specialize for each specific location or region. This could be useful especially since the results indicated that some sets of predictors may show a geographic bias (Fig. 8). However, using these geographic predictors would cost two vertical coordinate, moisture, or dynamic type predictors in a four predictor configuration.

Another possible option is to include more than four predictors in the LUT. However, since the total number of cells is the product of the number of categories for each predictor, adding predictors adds another term to that product. Currently every predictor has 51 categories (0-50), and there is four predictors, so that means there are 6,765,201 cells in a LUT. Adding one more predictor with 51 categories would increase that number to 345,025,251 cells. Clearly this growth can be limited by choosing a smaller number of categories for each or some predictors. Though no extensive research has been conducted into the effects of reducing the number of categories, it is safe to assume that the SCFM would suffer an extreme loss of accuracy if this were taken too far (e.g., 2 categories for each predictor). In fact, McDonald (2005) did notice a decrease in accuracy in the operational SCFM when he tried to decrease the number of categories for some predictors.

The impact of increasing the number of cells in the LUT is twofold. First, memory and storage limitations place a technical limitation on LUT size. Second, the larger a LUT becomes, the more observations would be required to get a significant sampling of cloud in each cell. This, in turn, would increase the amount of time and training data as dramatically as the number of cells increases when adding another predictor to the model. However, adding more predictors may improve the accuracy of the model once it is trained, so a balance between the number of categories per predictor and the number of predictors should be a goal for optimizing the accuracy of the SCFM.

Other than the practical and technical limitations of adding more predictors to the SCFM, increasing the number of predictors in the SCFM will only give a limited

improvement to the model. There is only so much information about the atmosphere available, and adding more predictors may just add the same information in another form. For instance, including divergence as a predictor would most likely not yield a large improvement for the SCFM-R since it is so closely related to vertical velocity, which is already a predictor. Another example is the vertical coordinate type predictors. Pressure, geopotential, potential temperature, and temperature are all highly correlated with each other, (Table 3). However, adding predictors with the same information in a different form can have very positive effects. For example in both DS1 and DS2, relative humidity reexpressed as relative humidity to the fourth power was among the top performers more frequently than relative humidity, but both relative humidity and relative humidity to the fourth power were never together in a top performing set. This shows that how predictors are expressed may be more important than how many predictors are used.

D. VERTICAL COORDINATES AND THE BOUNDARY LAYER

It was shown in Chapter V that the difference of boundary layer between the operational SCFM and the SCFM-R had a large impact, especially on the bias. In fact, the boundary layer could be represented much better. First, in the SCFM-R, the boundary layer cloud is treated as the mean of the 1000 mb to 925 mb layers. While the operational SCFM uses the 500 m AGL level for the boundary layer, it still uses the 900 mb layer in the calculation of total cloud, as does the SCFM-R. For most of the globe this should pose no problem, however, there should be major concerns for this approach in mountainous regions. In many mountainous regions, the surface pressure is below 900 mb, so some of the layers that the models are using for diagnosing cloud cover in the boundary layer are actually underground. This means that the SCFM can actually diagnose cloud underground. For the output of the NWP models, the values of the bulk meteorological parameters are extrapolated to obtain values on these levels, but they do not represent any real physical state in the atmosphere. For these reasons, the vertical coordinate in the SCFM should not be pressure levels. Any vertical coordinate that does not intersect the ground is preferable. However, GFS data is output on many pressure levels, making pressure levels convenient. Since geopotential is also typically output, data from pressure levels should be used in combination with geopotential height and

topography to interpolate values to height above ground level. Since aircraft altimeters typically use barometric pressure, and upper level flight levels are along pressure surfaces, it may be useful to also have information on pressure surfaces. For this reason it would be advantageous to output cloud cover information on pressure levels for the upper levels, and on constant height AGL surfaces at low levels. Either way, the total cloud cover should be calculated using levels that do not intersect the surface.

In the boundary layer, heat fluxes, moisture fluxes, and friction are much stronger and related more to model physics than dynamics. Additionally, vorticity and vertical velocity tend to be much weaker at the surface, at least at global NWP model scales. Over large bodies of water, such as the ocean, moisture is much more available, and over land cloud condensation nuclei are much more abundant than over the ocean (Rogers and Yau, 1989). This all suggests that the boundary layer may be treated more accurately by using different predictors than in the upper atmosphere. For instance, using lapse rate to detect the presence of an inversion with a land water/mask could detect the presence of a marine layer and low stratus. Further, the land/water mask could help account for differences in cloud condensation nuclei over continents and oceans as well as differences in moisture availability. Shear could be used to detect a well mixed boundary layer, and the dot product of the wind vector with the gradient of the terrain could be used to detect upslope and down slope winds. For these reasons, the boundary layer should be treated separately from the rest of the atmosphere when diagnosing cloud cover.

E. DISCUSSION OF VERIFICATION

As was discussed in Chapter III, the training and verification data for the SCFM-R comes from the WWMCA. The SCFM, and SCFM-R, only use the total cloud cover field for training and verification, however, the SCFM is by design a three dimensional model. So the SCFM, a three dimensional model, only verifies with the total cloud cover, which is two dimensional data. Unfortunately, there does not exist a verified, tested, and proven three dimensional cloud analysis to verify the SCFM. Therefore, there is currently no information about how well predictor sets verify for the vertical distribution of cloud. However, there is abundant data for cloud ceilings and the cloud cover from lower decks in the form of hourly surface observations, so perhaps the first

step to verifying the vertical distribution of clouds is to verify the SCFM at specific locations against this data.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS AND RECOMENDATIONS

A. CONCLUSIONS

From the results in Chapter IV, it is apparent that the top performing sets of predictors require predictors from all three types: the vertical coordinate type, the moisture type, and the dynamic type. Two diagnosis schemes, DS1 and DS2 were used; in DS1, if the LUT could not be used to diagnose cloud for a point because it did not have enough observations, then 0% cloud cover was diagnosed, and in DS2 the initial condition was used. Furthermore in DS2, points that corresponded to LUT cells which described cloud cover distributions that were too bimodal were also diagnosed with the initial condition. DS1 and DS2 both performed similar, with the exception that DS2 had generally about half the bias of DS1. Overall, temperature frequently was not among the sets of top performing predictors, while geopotential, pressure, relative humidity, relative humidity to the fourth power, cloud water, vorticity, and vertical velocity all were.

When comparing the SCFM-R to the operational SCFM, the major difference was the boundary layer, and the much better correlation and bias of the operational SCFM indicate that the operational SCFM handles the boundary layer much better. In Chapter V, the boundary layer and vertical coordinate system were discussed in more detail, and it was concluded that improvements to the model boundary layer should be made due to the different physical processes occurring in the boundary layer. Such improvements would mainly consist of using different predictors in the boundary layer, but also using height above ground level coordinates.

When comparing images of cloud diagnoses, it became apparent that even though vorticity was popular among the statistically top performing sets of predictors, there was a major problem with using it. Horizontal bands (Fig. 9) were present in images which vorticity was a predictor, whether it was absolute vorticity or relative vorticity. Despite this artifact, vorticity did verify well, and otherwise looked very realistic in the images. Currently it is not known why this horizontal banding occurred. Also, when inspecting an image where the predictors were pressure, temperature, relative humidity, and vertical

velocity (Fig. 8), it is apparent that these predictors had a geographic bias to under diagnose cloud in equatorial regions.

The initial condition used for starting without any entries in the LUT is unphysical. The critical relative humidities for condensation or deposition to occur are much too high at low temperatures and they are much to low at high temperatures. This has occurred because the difference saturation vapor pressure over water and ice was not taken into account when matching the cumulative distribution functions for relative humidity and ice. Since the initial diagnoses determine how cloud cover is partitioned in the vertical, the SCFM is starting its training with an unrealistic distribution of cloud in the vertical.

B. RECOMMENDATIONS

The simplest and most immediate changes that AFWA can implement on the operational SCFM are changing the predictors. Currently, AFWA should consider replacing temperature with another moisture predictor, such as cloud water, or another dynamic predictor such as relative vorticity. However, the cause of the horizontal banding when vorticity is used should be isolated and eliminated before it is used. Also, improvement may be seen in the operational model by reexpressing the relative humidity as relative humidity to some positive power (e.g., RH^4).

For more long term improvements, research should be conducted into how to better diagnose cloud in the boundary layer. Experimenting with different vertical coordinates and different predictors in the boundary layer are both recommended. Perhaps the greatest improvement can come from adjusting the initial condition to take the difference of saturation vapor pressures over water and ice into account. Since this is how cloud is partitioned in the vertical initially, it is important to come as close to a physically consistent model initially, so the SCFM can refine and improve upon it.

LIST OF REFERENCES

GIBBS: Global ISCCP B1 Browse System. 2006: (Available online at http://www.ncdc.noaa.gov/oa/rsad/gibbs/gibbs.html).

Bluestein, H. B., 1993: Synoptic-Dynamic Meteorology in Midlatitudes. Vol. II, Oxford University Press, 594 pp.

Campana, K., P. and Chaplan, ed, 2005: Technical Procedures Bulletin for the T382 Global Forecast System. (Available online at http://www.emc.ncep.noaa.gov/gc_wmb/Documentation/TPBoct05/T382.TPB.FINAL.htm).

Carlson, T. N., 1998: Mid-Latitude Weather Systems. American Meteorological Society, 507 pp.

Carter, G. M., H. R. Glahn, 1976: Objective prediction of cloud amount based on model output statistics. Monthly Weather Review, Boston, 104, 1565-1572.

Cloud Models Team, HQ AFWA/DNXM, 2005: A Comparison of Cloud Forecast Models: Advect Cloud (ADVCLD), GFS Cloud (GFSCLD), Stochastic Cloud Forecast Model (SCFM).

Crum, Timothy D. ed., 1987: AFGWC Cloud Forecast Models. AFGWC/TN-87/001.

Djuric, D., 1994: Weather Analysis. Prentice Hall, 304 pp.

Fujita, T. T., 1986: Mesoscale Classifications: Their History and Their Application to Forecasting. Mesoscale Meteorology and Forecasting, P. Ray, Ed., American Meteorological Society.

Glahn, H. R., D. A. Lowry, 1972: The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. 11, 1203. Journal of Applied Meteorology: Vol. 11, No. 8, pp. 1203–1211.

Hamill, T. M., T. Nehrkorn, 1993: A short-term cloud forecast scheme using cross correlations. Weather and Forecasting, Boston, MA, 8, 401-411.

Hoke, Dr. James E., Hayes, Capt John L., and Renninger, 2Lt Larry G., 1981: Map Projections and Grid Systems for Meterological Applications. AFGWC/TN - 79/003.

HQ AFWA/DNXM, 2005: Algorithm Description for the Cloud Depiction and Forecast System II. Offutt AFB, NE 68113.

Inoue, T., H. Kamahori, 2001: Statistical Relationship between ISCCP Cloud Type and Vertical Relative Humidity Profile. J.Meteorol.Soc.Japan, 79, 1243-1256.

Jensenius, J. S.,Jr, 1988: Objectively forecasting sunshine. Weather and Forecasting, Boston, 3, 5-17.

Kalnay, E., 2003: Atmospheric Modeling, Data Assimalation, and Predictability. Cambridge University Press, 341 pp.

Kidder, S. Q., Vonder Haar, Thomas H., 1995: Satellite Meteorology, An Introduction. Academic Press, 466 pp.

Kvamsto, N. G., 1991: An investigation of diagnostic relations between stratiform fractional cloud cover and other meteorological parameters in numerical weather prediction models. Journal of Applied Meteorology, Boston, MA, 30, 200-216.

--1993: An investigation of the cumulus cloudiness parameterization in northerly flows in the Norwegian Sea. Monthly Weather Review, Boston, MA, 121, 1434-1449.

McDonald, D., 2005: Personal Communication.

McDonald, D. G., 2003: 16th Mesh Stochasitc Cloud Forecast Model.

Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research, 2002: PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3.

Mocko, D. M., W. R. Cotton, 1995: Evaluation of fractional cloudiness parameterizations for use in a mesoscale model. Journal of the Atmospheric Sciences, Boston, MA, 52, 2884-2901.

Norquist, D. C., 2000: Cloud Predictions Diagnosed from Global Weather Model Forecasts. Mon.Weather Rev., 128, 3538-3555.

Norquist, D. C., 1999: Cloud predictions diagnosed from mesoscale weather model forecasts. Monthly Weather Review, Boston, MA, 127, 2465-2483.

Norquist, D. C., D. L. Aiken, and D. A. DeBenedictus, 1997: Cloud cover predictions diagnosed from global numerical weather prediction model forecasts. Phillips Lab Hanscom AFB, MA. (Available online at http://stinet.dtic.mil/).

Norquist, D. C., H. S. Muench, D. L. Aiken, and D. C. Hahn, 1994: Diagnosing cloudiness from global numerical weather prediction model forecasts. Phillips Lab Hanscom AFB, MA. (Available online at http://stinet.dtic.mil/).

Pukall, B. D., 1998: A Comparison of Advect Cloud Model and Fifth-Generation Mesoscale Model Total Fraction Cloud Forecasts. Master's Thesis, Air Force Institute of Technology. (Available online at http://stinet.dtic.mil/).

Rogers, R. R., M. K. Yau, 1989: A Short Course in Cloud Physics. 3rd ed. International Series in Natural Philosophy, Butterworth Heineman, 290 pp.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and Cloud Parameterization Studies with a Mesoscale NWP model. Monthly Weather Review, 117, 1641-1641-1657.

Weiss, M., 2001: AVN-based MOS ceiling height and total sky cover guidance for the contiguous United States, Alaska, Hawaii, and Puerto Rico. NWS Technical Procedures Bulletin No. 483, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 22 pp.

Whiton, R. C., E. M. Berecek, and J. G. Sladen, 1981: Cloud forecast simulation model. United States.Air Force Environmental Technical Applications Center, Technical Note, 126.

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences. International Geophysics Series, Academic Press, 464 pp.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Air Force Weather Technical Library
   Asheville, North Carolina

4. Air Force Institute of Technology
   Wright-Patterson Air Force Base, Ohio

5. Professor Philip A. Durkee
   Naval Postgraduate School
   Monterey, California

6. Lieutenant Colonel Karl D. Pfeiffer
   Naval Postgraduate School
   Monterey, California

7. Professor Carlyle H. Wash
   Naval Postgraduate School
   Monterey, California